THE MANCHESTER SCHOOL OF ECONOMIC AND SOCIAL STUDIES

VOLUME XXIX

Edited by

Professor B. R. Williams

1961

Reprinted by permission of the original publishers

KRAUS REPRINT LIMITED
Nendeln/Liechtenstein
1967

Printed in Germany

Lessing-Druckerei Wiesbaden

CONTENTS

No. 1. January, 1961.

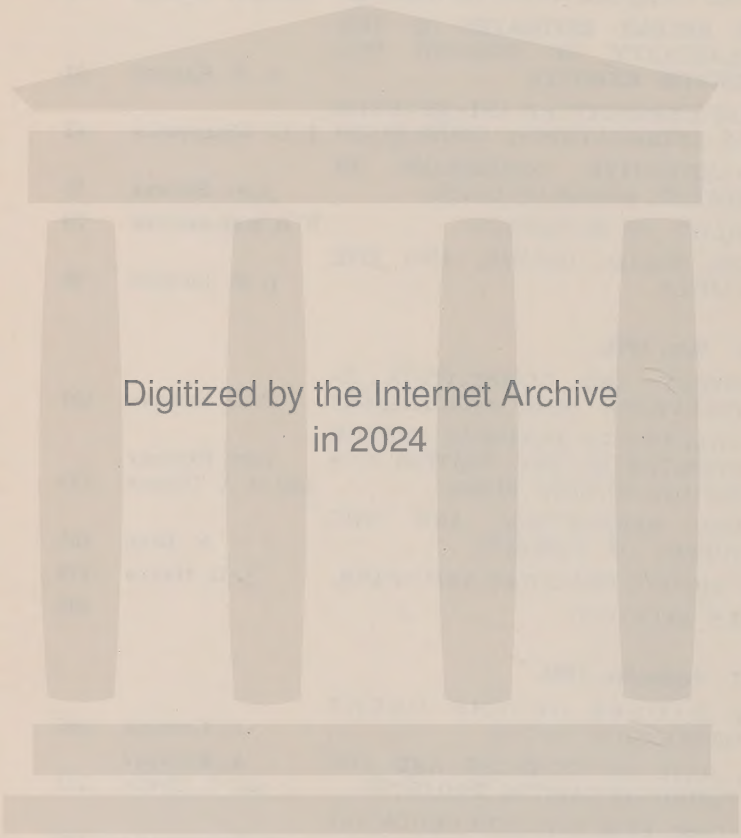
		Page
INTERNATIONAL FACTOR MOVEMENTS AND UNEQUAL RATES OF GROWTH	BRINLEY THOMAS	1
SOME RECENT ESTIMATES OF THE ELASTICITY OF DEMAND FOR BRITISH EXPORTS	S. F. KALISKI	23
PUBLIC EXPENDITURE AND REVENUE AN INTERNATIONAL COMPARISON	J. G. WILLIAMSON	43
ADMINISTRATIVE OVERHEADS IN BRITISH MANUFACTURING ..	JOHN BONNER	57
STABILITY IN OLIGOPOLY	R. B. HEFLEBOWER	79
WAGES, SOCIAL INCOME, AND THE FAMILY	J. M. JACKSON	95

No. 2. May, 1961.

MONOPOLY AND COMPETITION IN TELEVISION : SOME POLICY ISSUES	P. O. STEINER	107
AN ANALYSIS OF POST-WAR LABOUR DISPUTES IN THE BRITISH CAR MANUFACTURING FIRMS	JOHN BESCOBY and H. A. TURNER	133
CREDIT RESTRICTION AND THE SUPPLY OF EXPORTS	R. BALL	161
RESTRICTIVE PRACTICES AND AFTER.	J. B. HEATH	173
BOOKS RECEIVED		203

No. 3. September, 1961.

THE CAUSES OF THE GREAT DEPRESSION, 1873-96	D. COPPOCK	205
THE RATE OF INTEREST AND THE TIMING OF CAPITAL PROJECTS ..	A. MERRETT and A. SYKES	233
THE COST FUNCTION FOR CRUDE OIL PRODUCTION	G. BIEDERMANN	249
THE MEASUREMENT OF QUALITY CHANGES IN CONSUMER DURABLES	M. L. BURSTEIN	267
NOTES ON COST AND CAPACITY ..	BRUCE WILLIAMS	281



Digitized by the Internet Archive
in 2024

https://archive.org/details/manchester-school_1961_29

International Factor Movements and Unequal Rates of Growth[1]

The subject of this paper falls under the general heading of the relation between international factor movements and economic development. It discusses the effect of international movements of labour and capital since the second World War on the widening gap between rates of economic growth in advanced and under-developed countries. Some attention will be devoted to the migration of skilled labour and highly qualified personnel

Scale and Direction of Migration.

By migration we mean movements involving change of permanent residence from one country to another which take place through the volition of the individuals or families concerned. In this sense there has been a marked decline in international mobility as compared with the pre-1913 period. In the early years of this century the United States was receiving immigrants at an annual rate of 11 per 1,000 of her 1910 population : after the second World War the annual rate of inflow was only 1·3 per 1,000 of the 1950 population. If, however, we include the forced transfers of people from country to country which have been such a grim feature of our time, the picture becomes quite different. The partition of India and Pakistan uprooted no less than 18 million people, Western Germany absorbed 12 million refugees, Japan 6·3 million, South Korea 4 million, Hong Kong 1·3 million, Israel 1 million and Arab refugees from Palestine numbered about 1 million. The International Refugee Organisation and the Intergovernmental Committee for European Migration settled 1·3 million refugees overseas. This is by no means a complete account ; it leaves out entirely what has been happening in the Soviet and Chinese orbit. But even the cases given here yield a formidable total of 45 million "political" migrants—an international transfer of

[1]Based on a paper read to a joint session of the Economics and Sociology sections of the British Association for the Advancement of Science at Cardiff on 6th September, 1960.

people in one decade equal to the entire trans-Atlantic outflow from Europe in the century ending in 1913.

These forced transfers have had profound economic and demographic consequences, as we shall find when we deal with Western Germany. This paper, however, is mainly concerned with "economic" as distinct from "political" migration. The scale of inter-continental movements between 1946 and 1957 is given in Table I.

TABLE I

WORLD POPULATION, BY CONTINENT, IN 1946 AND 1957, AND BALANCE OF INTER-CONTINENTAL MIGRATION IN THE INTERVENING PERIOD
(in millions)

Continent or Major Region	Population		Total Population Increase	Net Migration
	1946	1957		
Africa	185·0	225·0	+ 40·0	+ 0·5
America	300·0	381·0	+ 81·0	+ 4·4
Asia	1,302·0	1,556·0	+254·0	− 0·5 (1)
Europe	379·0	414·0	+ 35·0	− 5·4
Oceania	11·8	15·4	+ 3·6	+ 1·0
U.S.S.R.	175·0	204·0	+ 29·0	—

(1) Excluding movements towards Israel.

Source : I.L.O., *International Migration*, 1945–1957, Geneva, 1959, p. 304.

The net outflow from Europe in the period 1946–1957 was 1·4 per cent. of the 1946 population or less than two years' natural increase, whereas net immigration to America was 1·5 per cent. of the 1946 population or less than one year's natural increase. The relative impact was substantial in Oceania where the rate of inflow was 9·3 per cent. of the 1946 population or more than five years' natural increase.

The experience of individual countries in Europe has varied considerably. Table II divides 17 European countries into three groups—Southern and Central-Northern, each with a net loss by migration, and another Central-Northern group with a net gain.

TABLE II
POPULATION GROWTH AND NET MIGRATION FOR CERTAIN EUROPEAN COUNTRIES, 1946–57
(in thousands)

Group and Country	1946 Population	Natural Increase	Estimated Net Migration		
			In absolute figures	As % of 1946 Population	As % of Natural Increase
Southern : losing.					
Austria	7,000	296	— 330 (1)	— 4·7	—111·0
Greece	7,480 (4)	1,230	— 150 (3)	— 2·0	— 12·0
Italy	45,290	5,154	—1,800 (3)	— 4·0	— 35·0
Malta	290	78	— 60	—21·0	— 77·0
Portugal	8,120	1,219	— 400	— 4·9	— 33·0
Spain	27,010	3,524	— 480	— 1·8	— 14·0
Central-Northern : losing.					
Denmark	4,100	497	— 70	— 1·7	— 14·0
Finland	3,810	661	— 100	— 2·6	— 15·0
W. Germany (2) ...	43,950	3,132	—1,300 (1)	— 3·0	— 41·0
Ireland	2,960	316	— 380	—12·8	—120·0
Netherlands ...	9,420	1,930	— 125	— 1·3	— 6·5
Norway	3,120	421	— 40 (3)	— 1·3	— 9·5
United Kingdom...	49,220	3,165	— 350 (3)	— 0·7	— 11·0
Central-Northern : gaining.					
Belgium	8,370	481	+ 160	+ 1·9	+ 34·0
France	40,280	3,538	+ 350 (3)	+ 0·9	+ 10·0
Sweden	6,720	543	+ 180	+ 2·7	+ 33·0
Switzerland ...	4,470	434	+ 220 (3)	+ 4·9	+ 51·0

(1) Except refugees and emigrants to Israel.
(2) Except national refugees (other than those arriving through Austria), foreign refugees and emigrants to Israel.
(3) Margin of error greater than 5 per cent. in either direction possible, owing to incompleteness of migration statistics.
(4) 1949 figure.
Source : I.L.O., *op. cit.*, p. 305.

Southern Europe, comprising Austria, Greece, Italy, Malta, Portugal and Spain, had a net loss by migration of 3,220,000 between 1946 and 1957 ; a Central-Northern group, comprising Denmark, Finland, West Germany, Ireland, Netherlands, Norway and the United Kingdom, had a loss net of 2,365,000. The only countries to register a net gain by migration were Belgium, France, Sweden and Switzerland ; for the four together it amounted to 910,000. Prominent among the losers

were Ireland and Austria ; the net outflow from both these countries exceeded their natural increase over the period. Malta lost three-quarters of her natural increase. On the other side, Switzerland and Sweden showed a vigorous absorptive capacity ; the inflow into Switzerland was equal to half her natural increase and added 5 per cent. to her 1946 population. The greatest absolute net outflows took place from Italy (1,800,000) and West Germany (1,300,000). It is important to note, however, that Table II deals exclusively with "economic" migration ; if the huge number of refugees received into West Germany and Austria were taken into account, these countries would show a very substantial inward movement.

Figures for the traditional receiving countries overseas are given in Tables III and IV.

TABLE III

POPULATION GROWTH AND NET MIGRATION FOR EIGHT AMERICAN COUNTRIES, 1946–57.
(in thousands)

Country	1946 Population	Natural Increase	Estimated Net Migration		
			In absolute number (1)	As % of 1946 Population	As % of Natural Increase
Canada (2) ...	12,620	3,305	+ 1,000	+ 7·9	+ 30
United States	141,390	27,735	+ 2,200	+ 1·6	+ 8
Jamaica ...	1,300	348 (3)	— 90	— 6·9	— 23
Mexico ...	23,180	8,689	— 250	— 1·1	— 3
Argentina ...	15,650	3,358	+ 800	+ 5·1	+ 23
Brazil ...	47,310	16,000 (4)	+ 450	+ 1·0	+ 3
Uruguay ...	2,280	320 (5)	+ 110	+ 4·8	+ 34
Venezuela ...	4,390	1,812 (3)	+ 330	+ 7·5	+ 16

(1) Margin of error greater than 5 per cent. in either direction possible, owing to incompleteness of migration statistics.
(2) And Newfoundland.
(3) 1946–56 only.
(4) Approximation based on population estimates.
(5) Approximation based on partial data.
Source : I.L.O., *op. cit.*, p. 308.

TABLE IV
AUSTRALIA AND NEW ZEALAND : POPULATION GROWTH AND NET MIGRATION, 1946–57
(in thousands)

Country	1946 Population	Natural Increase	Estimated Net Migration		
			In absolute numbers	As % of 1946 Population	As % of Natural Increase
Australia ...	7,460	1,391	930	12·5	67
New Zealand	1,660 (1)	351	145	8·7	41

(1) European population only.
Source : I.L.O., *op. cit.*, p. 312.

The fastest growing population in the Americas is the Canadian which increased by at least 1 million (or about 8 per cent. of the 1946 figure) through net immigation between 1946 and 1957. Too much reliance cannot be placed on this estimate because of the immense difficulty of gauging the balance of migration between Canada and the United States. Venezuela had a net influx (mostly from Europe) at almost the same rate (7·5 per cent. of the 1946 population), but it was only 16 per cent. of the natural increase as against 30 per cent. in Canada. The net movement into Argentina, estimated at 800,000, was equal to a quarter of her natural increase. The United States received an estimated net influx of 2,200,000, or 1·6 per cent. of the 1946 population, while Brazil was taking in immigrants at an even lower rate, 1·0 per cent. of her 1946 population. On the debit side, Jamaica lost through emigration nearly a quarter of her natural increase, and Mexico had a net outflow of 250,000 almost wholly to the United States, the equivalent, however, of only 3 per cent. of her natural increase.

One of the highest rates of net immigration in the world is found in Australia, amounting to 12·5 per cent. of the 1946 population. In New Zealand the corresponding rate was 8·7 per cent. The population of Australia grew by 31 per cent. between 1946 and 1957, three-fifths of it natural increase and two-fifths immigation.

In view of the disturbed state of affairs in certain parts of Africa, special interest attaches to the recent course of European

migration into some of these countries. There have been some dramatic increases in white population. The number of Europeans in the three territories of the Federation of Rhodesia and Nyasaland rose by 161 per cent. (from 105,000 to 274,000) between 1946 and 1957, and over two-thirds of the increase (116,000) was accounted for by immigration. Between 1946 and 1956 the number of white settlers in Kenya went up by 130 per cent. (from 24,900 to 57,700) and in Tanganyika and Uganda by 200 per cent. (from 12,100 to 36,100). These relatively large inflows are having a considerable effect on the natural increase. The white population of the Belgian Congo rose from 34,000 to 102,000 in the decade under review, two-thirds of the increase being due to immigration. Similar rates have been experienced in Portuguese and French territories ; in Morocco the non-Moroccan element in the population grew from 305,000 to 460,000 between 1946 and 1955. Reliable estimates of the movement of indigenous peoples are not available. Something is known, however, about the growth in the number of Indians and Arabs in British East Africa.

TABLE V

POPULATION OF ASIAN ORIGIN IN BRITISH EAST AFRICA

(In thousands : estimate at mid-year)

Territory	Indian Population		Arab Population	
	1946	1956	1946	1956
Kenya	88·4	151·9	23·3	33·0
Tanganyika	42·0	76·4	10·6	16·9
Uganda	32·0	54·3	1·4	2·0

Source : East African Statistical Department, *Annual Report*, 1955–56 (Nairobi, 1957).

Between 1946 and 1956 the number of Indians increased by 72 per cent. in Kenya, 82 per cent. in Tanganyika and 70 per cent. in Uganda ; it is safe to say that net immigration was responsible for at least half of this growth. The Arab population is much smaller and its rate of increase more moderate.

THE 19TH CENTURY PROCESS.

As background we shall glance briefly at the mechanism of international mobility in the 19th century. It can be seen most clearly if we forget national boundaries and regard the Atlantic community of nations as one single economy. This economy could not attain its warranted rate of growth without a redistribution of labour and capital as between its component parts. The flow of capital was mainly from Great Britain, whereas the whole of Europe contributed to the westward transfer of labour. The movement of factors did not proceed at an even pace : on the contrary there was a long swing with an average span of 14 to 20 years in trans-Atlantic migration of labour and capital, and an inverse relation between long swings in capital construction in the United Kingdom and in the factor-receiving countries overseas. The explanation of the inverse relation is to be found in the international movement of factors, particularly of population.[1]

Between 1821 and 1932 at least 60 million people crossed the ocean, the vast majority of them Europeans, 90 per cent. of whom settled in North or South America. Each of the great waves of emigration was for the most part a rural exodus, culminating in the greatest outflow of all between 1903 and 1913 when 10 million left Europe for the United States alone, 70 per cent. of them from the overpopulated areas of Italy, Russia, Central Europe and the Baltic states. The pattern of capital movement is shown by the fact that by 1913 aggregate British foreign investment amounted to £3,763 million distributed geographically as follows : 47 per cent. in the British Empire, 20 per cent. in the United States, 20 per cent. in Latin America, and 6 per cent. in Europe.

Four characteristics of the 19th century process may be briefly noted. First, factor mobility and inverse long swings in capital formation were a major reason why the international monetary system worked reasonably well. Secondly, the inverse long investment cycles made it possible for the effects

[1]For a detailed analysis of this mechanism see Brinley Thomas, *Migration and Economic Growth* (Cambridge University Press, 1954), chaps. VII, X, XI and XIV ; and "Migration and International Investment" in Thomas, ed., *Economics of International Migration* (Macmillan, 1958), pp. 5-11.

of differential productivity increases to be transmitted internationally without serious disequilibrium. Thirdly, the rate of economic growth in the factor-receiving countries was higher than in the leading creditor country, as shown in Table VI.

TABLE VI

RATES OF ECONOMIC GROWTH IN THE FIVE DECADES ENDING IN 1913

Country	Period	Percentage Change in Real Product per capita per Decade
United Kingdom ...	1860–9 to 1905–14	12·5
Germany	1860–9 to 1905–14	21·6
Canada	1870–9 to 1905–14	24·7
Sweden	1861–8 to 1904–13	26·2
United States ...	1869–78 to 1904–13	27·5

Source: S. Kuznets, "Quantitative Aspects of the Economic Growth of Nations, I, Levels and Variability of Rates of Growth," in *Economic Development and Cultural Change*, University of Chicago, Vol. 5, No. 1, Oct. 1956, p. 23.

In the half-century ending in 1913 the average increase in real product per capita each decade in the United States and Canada was twice as high as in the United Kingdom, and the same was also true of two European countries—Sweden and Germany—which experienced their industrial revolution late in the century. Under this regime it was not difficult for the less mature countries to adjust themselves to changes in the centre country. Fourthly, the process entailed a diffusion of economic power. When the United Kingdom was at her zenith in 1870 she was producing nearly one-third of the world's manufacturing output: by the early 'eighties the United States had overtaken her, and by 1906–10 Germany had done the same.

Another way of looking at 19th century experience is from the standpoint of the theory of factor price equalization. The highly restrictive assumptions of the Lerner and Samuelson models are not appropriate ; but on more reasonable postulates it can be demonstrated that free trade must be accompanied

by international migration of factors if total production in the international economy is to be maximized. As Professor Meade has shown,[1] we must allow for differences in the atmosphere in which production takes place in different countries as well as variations in scale ; we must also admit costs of transport, the existence of numerous factors relatively to the number of traded products, complete specialisation by certain countries, and substantial differences in the technical possibility of substituting one factor for another in different industries producing traded products. A model on these assumptions demonstrates the rationale of international migration.

Thus, both the analysis of long cycles in investment and the comparative-statics approach based on relative factor proportions provide an explanation of the part played by international factor movements in the 19th century in helping to narrow the gap between advanced and less developed countries.

Trans-Atlantic Capital Movements, 1945–60.

The period 1945–60 divides itself into two phases with a turning point about 1952. There was no longer any direct parallelism between migration movements and private foreign lending. In its place there emerged in the years 1945–52 a triangular mechanism based on the massive flow of *public* capital from America to Europe. In the immediate post-war phase the net transfer of governmental loans and grants from the United States to Western Europe was no less than $22,800 million. This was five times larger than the entire net amount of $4,400 million of American direct investment (i.e. investments by United States companies in their subsidiaries and branches abroad) throughout the world in the same period. The European Recovery Programme enabled the exhausted countries of Western Europe not only to cover their adverse trade balances but also to export capital (public and private) as well as migrants to the overseas countries within their orbit. Thus, straight after the end of the war, Great Britain, France, the Netherlands and Belgium were able to operate a regime of factor mobility and a wide range of multilateral trade. The

[1] J. E. Meade, *Trade and Welfare*, (Oxford University Press, 1955), Part III.

American private investor turned his back on Western Europe (only 8 per cent. of the total net direct investment in 1946–52 went to Western Europe) ; but his Government stepped into the breach and created a foundation on which the countries of Western Europe could rebuild safely. Meanwhile, the flow of American private capital to Canada and Latin America served indirectly to sustain the resumption of European emigration to those countries, and the Outer Sterling Area and the oversea dependencies of European countries were enabled to recover quickly. Such was the triangular mechanism which in the early post-war years contributed so much to the rapid repair of the economic damage done by the war.

The second phase which began about 1952 has seen a remarkable economic resurgence in Western Europe. Between 1952 and 1959 industrial production in the O.E.E.C. group of countries increased by 50 per cent. (Western Germany 78 per cent., Italy 73 per cent., France 55 per cent., and the United Kingdom 29 per cent.), while the corresponding rate in the United States was 20 per cent. Gross national product per head rose by 32 per cent. in O.E.E.C. countries, 46 per cent. in Western Germany, 21 per cent. in the United Kingdom and 7 per cent. in the United States. One is reminded of the 19th century pattern when Great Britain was the leading creditor ; phases of vigorous British investment in oversea countries were followed by phases in which the production and export capacities of these countries showed a rapid increase. The extraordinary upsurge in production and export capacity in Western Europe in 1952–60 is a corollary of the basic capital formation which American aid in 1945–52 made possible.

The relatively rapid growth of productivity in Europe had the expected effect on the balance of payments of the United States. There was a redistribution of reserves ; between 1952 and 1960 the official gold and foreign exchange holdings of Continental Europe rose from just under $9,000 million to $20,000 million and those of the United Kingdom from $2,000 million to $4,000 million, whereas in the United States there was a decline from over $23,000 million to $18,000 million. The outstanding facts about United States factor movements in this period were the continuance of large exports of public

capital in the form of military expenditure and grants and the remarkable revival of private foreign investment mainly by subsidiaries and branches of American firms abroad. In 1959 for the first time the volume of new American funds for direct investment in Western Europe was greater than in either Canada or Latin America. The total outflow of long-term public and private capital from the United States in 1960 was $8,500 million and in addition there was a transfer of about $2,000 million of short-term assets ; this entailed a balance of payments deficit of $4,000 million. Towards the end of the year active measures were being taken to stem the gold drain.

The deficit was not due to any weakness in the American economy : it was the result of the enormous scale of America's defence expenditure abroad at the annual rate of $6,000 million. If this could have been avoided, the growth in *private* foreign investment particularly to Western Europe in the years 1952–60 would have gone far towards ensuring the stability of the triangular mechanism which began with the flow of *public* capital exports for European recovery in the years 1945–52. Another important condition, however, would have to be fulfilled, namely, that surpluses piling up in continental Western Europe, especially Western Germany, would be put back into world circulation through an appropriate increase in imports and investment in under-developed countries.

THE RANGE OF INEQUALITY.

Professor Kuznets, in his path-breaking investigations into the quantitative aspects of the economic growth of nations, divided the world (in descending order of 1949 relative income per capita, world = 100) into the following groups : overseas descendants of Western Europe (i.e. U.S.A., Canada, Australia and New Zealand) (590), Western Europe (214), U.S.S.R. (133), Other Europe (94), Latin America (66), Africa (24), and Asia (20). He has produced strong evidence in favour of the following propositions.

(*a*) Over the half-century 1900–1950 the growth of income per capita in group I (U.S.A., Canada, Australia and New Zealand) taken together was well over 100 per

cent. and in group 2 (Western Europe) well over 50 per cent. and perhaps nearer 75 per cent.

(*b*) "The weight is quite heavily in favor of the assumption that the rates of growth during the last half-century in per capita income in the "rest of the world" must have been much lower than those established for the advanced countries of regions 1 and 2. And there is little, if any, evidence to the contrary. The disparity between the advanced countries and the "rest of the world" in per capital income, must then, have increased since 1900." [1]

(*c*) The countries which are now in the two top income brackets were well ahead of the "rest of the world" at the very outset of the process of modern industrialization. What the latter has done is to increase the original disparity. "While the presently developed countries entered the industrialization process from the position of advance, or superiority, at least with respect to per capita income, to the "rest of the world," the currently under-developed countries in the "rest of the world" face the prospect of industrialization from a position of inferiority in that their per capita incomes lag substantially behind those of the advanced sector of the world's economy to-day." [2]

MOBILITY AND INEQUALITY SINCE 1945.

1. *Factor mobility a cause of biased growth in the rich sector.*

Four oversea descendants of Western Europe—the United States, Canada, Australia, and New Zealand—now form the world's top income bracket, and in recent years Western Europe itself has been taking long strides towards their standard. This aristocratic club of industrially advanced economies has an average standard of living enormously higher than that of members of the four lowest income brackets—Other Europe, Latin America, Africa and Asia.

[1] S. Kuznets, "Quantitative Aspects of the Economic Growth of Nations. I, Levels and Variability of Rates and Growth," *Economic Development and Cultural Change*, Chicago, V, 1, October, 1956, p. 23.

[2] *Ibid.*, p. 25.

We have observed one interesting similarity between present tendencies and the 19th century process. The rate of growth in the factor-receiving countries exceeds that of the capital-providing country. The gap between the United States and Western Europe in the last fifteen years has been narrowing ; the same thing happened as between Great Britain and the oversea receiving countries in the 19th century. However, this apparent similarity is overshadowed by real differences. The main factor-receiving countries are now the richest in the world, and the peculiar pattern of capital movements in the Atlantic community since 1945 has given an immense advantage to Western Europe. The circulation of public and private capital and of labour within this affluent sector tends to increase the disparity between it and the under-developed areas, whereas in the 19th century the transfer of labour and private capital from Europe to oversea countries enabled a large part of the then under-developed sector of the world to achieve a relatively rapid rate of growth.

Tables I, III and IV above show that in the period 1946–57 total net inter-continental migration was 5,900,000. The four countries with the highest standard of living, United States, Canada, Australia and New Zealand absorbed 4,275,000, or 73 per cent., of this total. These are the classic immigration countries, and it is not surprising that they should continue to be attractive to European settlers.

Trends in capital movements since the war show that of the £2,425 million absorbed by the Outer Sterling Area, 72 per cent. went to Australia, New Zealand, South Africa, and Rhodesia, and 28 per cent. to India, Pakistan and the colonies. Of the £1,225 million of public capital, 40 per cent. went to Australia, New Zealand, South Africa, and Rhodesia, and 60 per cent. to India, Pakistan, and the colonies. The £500 million of private capital invested in the colonies included £175 million for petroleum (£50 million in Aden, £100 million in Trinidad, and £25 million in Nigeria) and £50 million for bauxite in Jamaica. Only 18 per cent. of the total private investment in sterling countries went into investments other than oil and bauxite in the least developed parts of the Commonwealth and Empire. Mr. A. R. Conan, whose valuable researches are the basis of

the above figures, reaches the following conclusion. "It is apparent . . . that the post-war flow of capital has not in the main been towards those countries which most urgently needed development resources. On the contrary, the trend has been perverse to the extent that the available supplies of investment capital have gone mainly to countries which had already attained a relatively high degree of development." [1]

Recent figures on American direct investments abroad reveal similar tendencies.[2] The aggregate value of United States direct investments in 1959 was $29,735 million, of which $10,423 million, or 35 per cent., was in petroleum. The distribution over various parts of the world was as follows.

	Canada	Europe	Latin America	Africa	Asia	Oceania	International
	%	%	%	%	%	%	%
Total U.S. Direct Investments (excluding petroleum)	40	20	28	3	3	3	3
U.S. Direct Investments in Petroleum ...	24	14	32	3	16	3	8
U.S. Direct Investments in Manufacturing ...	47	30	16	1	2	4	—

The bulk of the foreign operations of American companies —60 per cent. of total investments other than petroleum and 77 per cent. of manufacturing—is in the highly developed North Atlantic economy : the proportion in Asia and Africa is negligible, i.e., 3 per cent. or under. It is only in petroleum that the balance is tilted slightly in favour of the underdeveloped countries, 32 per cent. of these investments being in Latin America, 16 per cent. in Asia, and 3 per cent. in Africa. However, the trend in the last three years has been markedly in favour of the more advanced countries. The development of oil resources in the Middle East and some parts of Latin

[1] A. R. Conan, *Capital Imports into Sterling Countries*, Macmillan, 1960, p. 37.

[2] See *Survey of Current Business*, Washington, Sept., 1960, pp. 15-24.

America has slowed down appreciably ; half of the $300 million spent on mining properties in 1959 was incurred in Canada ; some major iron ore and other mining ventures along the West coast of Africa have been hindered by the unrest in those areas. Meanwhile, there has been a spectacular rise in the scale of American investment in manufacturing in Europe and Australia; between 1957 and 1959 the rate of increase was 38 per cent. in Europe and 31 per cent. in Oceania (mainly Australia) as against only 11 per cent. in Latin America.

So far we have restricted our attention to inter-continental movements of labour and capital. A full analysis would have to include the redistribution of factors across national boundaries within continents. Part of the picture for Europe is given in Table II above, but it has to be supplemented by taking into account the considerable transfers of refugees. The main countries of absorption are Western Germany, France, Switzerland, Sweden, and Belgium. The high rate of economic growth and the development of the European Common market have resulted in a large volume of internal migration within the continent, some of which has been at the expense of the flow of emigrants to oversea countries.

Western Germany presents one of the most striking examples of mass immigration on record. The population increased by one-third in 12 years. Between May, 1945 and July, 1959, Western Germany absorbed 3·2 million refugees from the Soviet Zone and 9·4 million from beyond the Oder-Neisse, from the Sudeten area and from South-eastern Europe. The labour force grew from 21·7 million in 1950 to 25·3 million in 1957. Although the circumstances were extraordinary, the process of absorption was reminiscent of the greatest inflows of migrant labour into the United States in the 19th century. The newcomers included a high proportion of skilled labour and highly qualified personnel ; they spoke the same language and had the same background as the people in the receiving country, and were highly adaptable and mobile ; the pressure in the labour market made the trade unions docile and kept down the level of real wages ; the receiving economy had an advanced industrial technique, and the huge input of labour concentrated in a short period induced a "widening" of the

capital structure with potent effects on the growth of productivity ; and the large inflow of capital in the form of aid together with the high rate of accumulation promoted by Government policy resulted in a substantial rate of capital formation. Intra-continental movements of labour in Europe have been to an increasing extent a substitute for European emigration, and this is making it more difficult for the poorer immigration countries overseas to satisfy their demands, particularly for the more skilled grades.

2. *"Perverse" movements of highly-skilled personnel.*

International movements of various grades of skilled labour and highly qualified personnel should be treated as part of the flow of capital. This species of capital is indispensable to the process of growth in an under-developed country, whether it is under- or over-populated. Where factor scarcities are such that immigration is a condition of development, the absorption of unskilled workers will be governed by the rate of inflow of complementary skilled grades. Where there is serious overpopulation with no possibilities of emigration, what is required is the import of a small nucleus of skilled labour and highly qualified personnel ; such a nucleus acts as a centre for the diffusion of skills. The supply of entrepreneurs, engineers, managers, administrators, teachers, accountants, tax-collectors, foremen, and skilled craftsmen is very small in relation to the vast need. It is more economical for a given under-developed area in the short run to import skill than to train local candidates ; if it succeeds, it can expect that in the long run the migrants will transmit their skill to the local population. Migrations of highly qualified personnel, even on a very small scale, between different Asian communities can have a considerable marginal social product, if they are accompanied by measures to ensure that difficulties of assimilation are eased and the channels for the spread of technical knowledge are kept open. The essential part played by the mechanism of diffusion of skills was brought out by Professor Lockwood in his study of the economic growth of Japan. "With all its limitations, both in scale and content, this system of general and technical education as it expanded was fundamental to

industrial advance. In time Japan came to have a ratio of engineering students to the population said to be higher than in many industrial nations of Europe. In pure science, still more than in engineering, her universities and research institutes gained the respect of the world. The mass of people never got beyond six years' schooling. But even this helped to diffuse the more modest technical skills widely among the population, especially in the towns and cities. And education encouraged an increasing degree of geographic, occupational and social mobility among the people which was essential to the radical restructuring of the economy . . . No other enterprise of the State paid handsomer dividends to the nation. The chief trouble was there was not enough." [1]

The trends in factor mobility in the last 15 years indicate an intense demand for skilled personnel in North America and Europe. Can the under-developed countries hope to compete with the more advanced for the limited supplies of skill? Is this disability a major cause of the widening gap between the rate of growth in the poorest nations and in the more mature economies? In contrast to the long-run community of interest between the investing countries and the emergent oversea areas of new settlement in the 19th century, is there now a deep and cumulative conflict of interest?

Unfortunately, there are big gaps in our empirical knowledge of international movements of skilled labour and highly qualified personnel. At this stage one can only form an impression based on incomplete data. Among the sources available special mention should be made of the investigations carried out by the O.E.E.C. (e.g. *Shortages and Surpluses of Highly-qualified Scientists and Engineers in Western Europe*, Paris, 1955) and a survey by the Economic Commission for Europe ("Highly qualified Technical Manpower in Europe," *Economic Bulletin for Europe* 9 (1), May, 1957, pp. 41–63). Another example of the kind of research needed is the I.L.O.'s *Report on International Migration of Labour in the Construction Industry* (I.L.O., Building, Civil Engineering and Public Works Committee, 6th Session, Geneva, 1959) which contains detailed

[1] William J. Lockwood, *The Economic Development of Japan* (Princeton University Press, 1954), pp. 511-512.

statistics of the movement of various grades of construction workers into and out of the United Kingdom, Australia and Argentina, and into the United States and Canada. The significance of this type of recruitment to a growing country may be seen from the fact that in 1957 one-third of the houses and buildings of all kinds under construction in Australia were being built by skilled immigrant tradesmen, and one-fifth of the workers employed on major public works construction projects were immigrants.

The annual rate of increase in the number of engineers in Western Europe, about 1 to 2 per cent., is well below the rate required to sustain development in the next ten years, even if there are no major technical innovations. Merely to increase the capacity of universities, technical colleges, and schools will not be enough : the training systems themselves will have to be made more flexible to ensure that the available talent comes within the field of recruitment. Scientists, engineers and other skilled personnel are attracted from Europe to the richest countries where salaries and amenities are attractive ; for example, many Austrian engineers have emigrated to the United States in the last fifteen years. Just as a large part of international trade is between industrialized nations, it is natural and beneficial that there should be a two-way traffic of skilled manpower.

What are we to say of the two-way traffic of skilled personnel between the poorest and the richest sectors ? On the one hand, we find a "perverse" emigration of trained people from the countries in the lowest income bracket to the advanced countries ; on the other hand, there is the reverse movement of experts financed out of public funds. Nearly every country in the Middle East is suffering from a considerable leakage of talent. Even in Israel, whose very existence is founded on immigration, the problem has become acute. In 1959 the number emigrating from Israel was 7,000 which was nearly one-third of the number coming in, and they were nearly all skilled workers and professional men. Mr. Nehru has called attention to the number of Indian scientists who have settled abroad, and the Government of Iraq has addressed appeals to

students who seem to use every resource in seeking to extend their period of residence in foreign universities.[1]

In a free society the individual has the right to come and go as he pleases. The owner of a scarce skill in a very poor country, even if he is offered the maximum income attainable at home, may well be led by net advantages to take employment in a rich country. Under a regime of *laisser faire laisser passer* the transfer of skilled personnel from poor to rich countries could be on such a scale that the former could never begin the process of growth. We must not be misled by some of the analogies drawn from the experience of under-developed countries in the 19th century. In that period, for reasons mentioned earlier in this paper, there was a natural flow of private capital and technical and entrepreneurial skills from the advanced country to the under-developed areas, and a great deal of the investment went into public utilities. If the under-developed countries of to-day, which are relatively much poorer and much more handicapped than their predecessors of a century or more ago, are to overcome their disabilities, measures must be taken to protect and increase their supplies of skill. It is important that public capital in the form of technical assistance programmes should be supplied by the rich countries, but foreign experts on short visits are an unsatisfactory substitute for an indigenous supply of trained personnel and a properly organised system of education.

3. *Instability of factor inflow based on non-competing groups.*

One of the striking features of the last fifteen years has been the remarkable influx of European capital and settlers into Africa. We have already noted the sharp increases in white immigration into the territories of East and West Africa. We can divide the countries south of the Sahara into two categories: (*a*) those which have attracted only European immigrants and

[1]The Shah, in a recent address to students going abroad, said, "Whether you like it or not, the future of the country depends on you . . . We and you should try to persuade those students who have gone before you, and who do not remember their country and homes, to return. Is Iran a backward country? Is it unorganized? Is it all bad? If they say it is good, they should come back to a good country. But if it is a bad country, then they should come back and help it to become better." (Quoted in *The Times*, 11th Oct., 1960).

are least developed, e.g. the French and Portuguese territories ; (*b*) those which have absorbed not only Europeans but unskilled African labourers to work in mines, industrial establishments, and agriculture, i.e. the Federation of Rhodesia and Nyasaland, Ghana, British East Africa, Congo, and the Union of South Africa. In some of these countries economic development has been vigorous enough to entail a demand for skilled African personnel which cannot be met internally.

The substantial inflow of European labour and capital into this part of the world is based on a system of stratification. There are three non-competing groups—European, Asian and African—meeting demands at three different levels ; there is no mobility between these three sectors divided on racial grounds. Here again is a glaring contrast with the 19th century process, in which European capital and labour went into under-developed oversea countries which had a very high degree of vertical social mobility. The economies of these African territories are founded on vertical immobility, and for this reason the future of European immigration is precarious. A system which does not allow adequate opportunities for Africans to receive training in skills and to obtain occupational and social advancement cannot long survive. If white capital and labour have gone to Africa on the assumption that the stratification will persist, these factor movements will not have contributed much to the progress of the under-developed world. It is now clear that the system is unstable. The appalling economic dislocation in the Belgian Congo is a grim object-lesson. The tripartite regime of non-competing groups based on racial discrimination is no longer viable.

Conclusion.

In the last fifteen years most of the international migration of capital and labour has been a circulatory process within the richest sector of the world, and this has helped to widen the gap between rates of growth inside it and in the poorest under-developed sector outside.

The disabilities of the under-developed areas are made worse by "perverse" movements of skilled labour and highly qualified personnel from these areas to the richer countries.

The remarkable influx of European labour and capital to Africa is based on a system of non-competing groups which does not allow opportunities for the African to be trained in skills and to obtain occupational advancement. This system together with the factor inflow associated with it is inherently unstable.

BRINLEY THOMAS.

Cardiff.

Some Recent Estimates of 'the' Elasticity of Demand for British Exports— An Appraisal and Reconciliation[1]

The post-war period has seen a great many attempts to estimate the parameters of various import and export functions.[2] In part, no doubt, this can be accounted for quite simply by increasing interest in the use of statistical data to quantify all sorts of economic models. One may regard Chang,[3] Neisser and Modigliani,[4] and Polak,[5] for instance, as trying principally to quantify the "Keynesian" system for a world economy, in much the same way as numerous research workers were simultaneously trying to quantify it for a number of national economics.

In part, however, and especially so far as the estimates of demand elasticity for exports, with which this article deals, are concerned, these attempts seem to have been stimulated by two problems peculiar to the theory of international economic policy. The first of these is the problem of exchange stability—of whether or not a country can improve her balance of payments by devaluing her currency. The second is the problem of commercial policy in the narrow sense—of the determination of the optimum rate of tariff or its equivalent in terms of other protective policies. Both these policy questions were, or were thought to be, of considerable urgency and importance in the post-war period. Moreover, in order to be able to offer advice even as to the *direction* which policy should take in both these instances—whether the external value of domestic currency or the level of the tariff should be raised or lowered—one requires fairly precise *quantitative* information concerning the relevant elasticities of demand and supply.

It is indicative of the sense of urgency with which economists felt the need for these policy recommendations that attempts at estimating foreign trade elasticities continued in

spite of the fact that these estimates were, and were constantly being shown to be, subject to very formidable statistical difficulties. Most of these difficulties were of a general nature, but the international trade estimates were often particularly subject to them.

The various points have been often and admirably stated elsewhere [6] and only the general conclusions can be related here. Briefly, the critics were able to show that all the estimates here discussed, and many others in this field, are lacking in scientific rigour. They are, that is, the results of estimating procedures the use of which, for this purpose, cannot, strictly speaking, be justified. The assumptions upon which these procedures are based are often not valid for the data and models to which they were applied. As a result, the estimates yielded are suspect on these technical grounds, quite apart from the question of the accuracy of the data on which they are based or the legitimacy of the inferences drawn from them.

Fortunately, the basic procedure involved, that of least squares, seems to be very sturdy, in the sense that quite serious violations of one or more of its assumptions only limit (often to an unknown degree) the accuracy of the results obtained and do not vitiate them completely. It is also the view of many authorities that least squares estimates, even when they are highly imperfect, are, in practice, often preferable to both such more scientific estimates as are obtainable and to no estimates at all.[7] To say this, however, is not to imply either that the statistical deficiencies of these least squares estimates are unimportant or that the estimates have always been interpreted with sufficient caution. Indeed, to the writer's mind, it was the chief positive contribution of the critics in this field to induce a more cautious attitude to such estimates.

To sum up the discussion so far, it appears that the international trade estimates were produced, in the face of great technical difficulties, because they were thought to be badly needed. The spirit of this work seems to have been that of A. J. Brown's suggestion that "further empirical work upon the elasticities of demand and supply and related coefficients in foreign trade, though exceedingly difficult, would be abundantly worth while if it achieved any results at all." [8]

It is in this spirit also, that some of this work will now be discussed. The question put, in other words, is not how far and in what ways the estimates fall short of ideal standards, but whether they offer any guide at all to the relevant magnitudes and how they might best be interpreted.

The focus of the discussion is narrowed at this point in yet another way. This article is not intended as a comprehensive survey of the literature of estimation in the field. Instead, some of the recent work of Marris,[9] Polak, Neisser and Modgliani, Chang, [10] and Zelder [11] and Harberger[12] is compared and examined in the light of some results obtained by the writer himself. These selections are intended to be broadly representative of recent published work having as an end product an estimate of the price elasticity of demand for global merchandise exports of the United Kingdom or large subdivisions of these exports. Marris's work is devoted entirely to this task, and to drawing policy conclusions from the results. The other studies are all broader in scope. All the estimates are based on interwar time series.

A summary of the numerical results is presented in the accompanying table. They will be seen to differ, and it has sometimes been suggested that they contradict one another. Some of the figures clearly refer to different groups of goods and are not, therefore, directly comparable. Although some of the remaining differences do, in fact, have to be explained merely in terms of differences in data and estimating procedures, some others are accounted for by the fact that the concepts of elasticity used by the several authors (explicitly or by inference from their estimating method) are radically different.

Indeed, to the writer's mind, the most striking thing about the table is the numerical similarity (and the direction of the divergence) of the first two (and latest) estimates. This is all the more remarkable because, to judge from their comments and policy recommendations, the author of the first of the estimates belongs to the "low elasticity" school of thought and the authors of the second to the "high elasticity" school.[13] Both seem confirmed in their convictions by their findings.

ESTIMATES OF PRICE ELASTICITY OF DEMAND FOR BRITISH EXPORTS
INTER-WAR PERIOD

Author	Group of Exports	Estimate	Notes
Marris	Total	1·8, 2·3	From Marris's linear equation $Y = -115 + 2{\cdot}15X$ and the extreme prices—those for 1933 and 1928–29—read off from his Chart II. The average elasticity for 1930–38 is 2·15.
Zelder-Harberger	Total (?)	at least 1·35	Harberger's estimate based on Zelder's elasticities of substitution R.E.S., p. 125.
Chang	Total	·53	p. 122.
Neisser-Modigliani	Manufactures to "primaries" countries	·66, ·80, ·72	For 1928, 1932 and 1935, respectively, from a linear equation, Table 17.
	Other	—	Estimates were attempted, but statistically significant coefficients with the right sign were not obtained. Neisser and Modigliani, p. 11, Polak, p. 72.
Polak	Manufactures	—	
	Non-manufactures	·92	Average for 1924–1938, p. 113.

All page numbers refer to works quoted. None of the authors gives errors of estimate.

There are also considerable conceptual differences between the two elasticities, which will be examined presently.

Before turning to a more detailed discussion of the estimates it may be useful to recall the outlines of the several authors' work :

Marris calculates the simple regression of U.K. exports on U.K. terms of trade, both expressed as index numbers 1930–1938 = 100, for the years 1920–1938, excluding 1926. He then transforms the relation into a foreign offer curve, discusses the interpretation of this curve, and examines how well it fits adjusted post-war data. His conclusion is that the curve, with some modifications, is still reasonably applicable. Marris suggests that the marginal cost of transforming exports into imports through foreign trade is shown by the curvature of this offer curve to be rather high. He argues that this evidence constitutes a *prima facie* case for a revision of policy with respect to foreign trade and protection.[14]

It is clear that the purely mathematical process of transforming the estimated relationship into a foreign offer curve, although it is a very effective expository device, can add nothing to the quality and meaning of the estimates. It will be found convenient, in what follows, to discuss the relationship between exports and terms of trade directly.

Polak fits a regression between British exports and world trade,[15] both expressed as index numbers, separately for primary commodities and manufactures for the years 1924–38. He finds that the relationship for primary commodities is improved by considering relative prices also, but that for manufactures is not. His estimating equation for manufactured exports, which also includes a time trend, indicates an "income" elasticity [16] of demand of ·75. For non-manufactured exports, the elasticity with respect to world trade in the same category is about a third and that with respect to relative prices about ·9, as indicated.

Having performed similar calculations for a number of other countries besides Britain, Polak finds relative prices significant in less than half his export equations.[17] He considers even those price coefficients which are statistically significant[17a] of doubtful value, partly because of the inappropriateness of his

relative price series, which are the ratios of a given country's export price index to that of the whole world's trade in the same category.[18]

Similar price relatives are used by Neisser and Modigliani.[19] They also find their price coefficients less reliable than the others [20] and only 10 of their 37 export equations improved by the inclusion of a relative price term.[21] Curiously enough in the light of Polak's results, one of these is the explanation of British exports of manufactured goods to "primaries" countries. For these the "typical" elasticity for the period is ·66–·80.[22] No elasticities are given for other British exports.

Neisser and Modigliani divide imports and exports into manufactures, foods, and raw materials. Their trading world consists of six members : the United Kingdom, the United States, Germany, France, the seven minor industrial countries,[23] and the primaries countries. The authors then estimate, by least squares, 36 structural equations, some of them subdivided by periods, explaining each country's import and export volumes in each category. The imports are explained in terms of industrial output, income,[24] food production and, sometimes, prices. The exports are expressed as functions of the several other countries' imports, sometimes combined, and occasionally, prices.

The price and industrial production variables are index numbers 1928 = 100, and the other variables in millions of 1928 dollars. All equations are linear in the variables.

The system is then solved, all imports and exports being, so far as possible, expressed as functions of the exogenous variables : incomes, prices, industrial production, and capital imports. Price and income elasticities for the years 1928, 1932, and 1935 are computed from the solved system.

Chang fits a log equation between the quantity of total U.K. exports, world real income, and a price relative which seems to be a ratio of the unit value index of total U.K. exports to an average of similar indices for the U.S.A., France, Germany, and Japan.[25] He finds an income elasticity of 1·8 and a price elasticity of ·5. Unfortunately all of Chang's variables are deviations from unspecified trends and it has been suggested

that the selection of these trends affects the value of his coefficients.[26]

A different attack on the problem of estimating the influences of relative prices was made by Zelder and Harberger. These writers despaired of calculating price elasticities directly from a regression of quantity exported on one or more determining variables.[27] Instead, they suggest the estimation of elasticities of substitution of particular commodity exports between two national sources by simple regressions of price relatives on quantity relatives or, *vice versa*. On certain assumptions, these elasticities can then be transformed into "partial" elasticities of demand with competitors' absolute prices, (and all other variables) held constant.

Zelder presents 37 such elasticities for the United Kingdom's inter-war exports in various commodity groups.[28] He suggests that these imply an elasticity of substitution for total exports of around 3.[29] This can be translated into an elasticity of demand of about 2·4.[30]

It was suggested above, that Marris's elasticity estimate is conceptually (and not merely computationally) different from the Zelder-Harberger estimate. It is, in fact, different from all the other estimates. A clue to this difference is provided by the fact that Marris relates the "quantity" of exports to terms of trade rather than to relative prices, but it goes much deeper than that :

Marris's estimate is, in principle, an estimate of *total* price elasticity while all the others refer to more or less *partial* elasticities.[31] The traditional theoretical concept of elasticity is clearly a partial (or *ceteris paribus*) one : It measures the effect of prices (or incomes) on the quantity sold, as expressed by a particular demand curve whose "shift parameters" are assumed constant for purposes of analysis. Needless to say, historical "shift parameters" do not normally behave in this accommodating fashion. The demand for British exports is affected by a whole host of variables other than their own prices or terms of trade. These include competitors' prices, the taste patterns of foreigners, their production possibilities and the degree to which these are utilized (often summarised as

"income"), barriers to trade, both natural and artificial, and the loans Britons make abroad.

Of course, movements in these shift parameters would not greatly matter if they were not associated with, or correlated to, movements in the price or terms of trade variable. Such "random" movements would, to be sure, reduce the precision with which one can estimate the effect of prices on quantities sold, but they would not affect the expected value or meaning of one's estimate. If the movements in the several explanatory variables are associated—and it requires little reflection to see that they are likely to be—then one is faced with the choice between considering the partial and the total relationship between price and quantity exported. The first of these would show the influence of price on quantity when all the other explanatory variables are held constant at given levels. The second would include, in addition, the full repercussions of price on the quantity demanded, indirectly, through the other explanatory variables.[32]

One would attempt to estimate the partial relationship by multiple regression techniques [33] which, in favourable circumstances (roughly, if the several explanatory variables move sufficiently independently of one another so that their net influences can be separated) allocate to each explanatory variable its partial influence. Clearly, no practical technique can consider all possible explanatory variables. Ultimately, the validity of the partial regression coefficients obtained depends upon that of the assumption of independence between the error term of the regression equation (which includes all residual influences) and the several explanatory variables. If this assumption is not justified, then one or more of the coefficients estimated is biased or, as an economist may prefer to put it, contains influences other than the partial influence of the variable to which it attaches.[34]

To estimate the total influence of price on the quantity demanded, one would fit the simple regression of the latter on the former. In favourable circumstances, the coefficient so obtained will approximate the total (direct and indirect) influence of price on quantity demanded. If the circumstances are not favourable, then the estimate has no particular claim

to be designated a measure of the influence of price. It then simply refers to the relation between the quantity demanded and a particular complex of historical events determining it.[35]

The notion of "favourable circumstances" is here more complex than in the case of multiple regression analysis. It involves economic as well as statistical judgments.[36]

Circumstances are favourable to the estimation of the total influence of, say, price, if there are no *autonomous net* movements of the other explanatory variables of a sufficient size to distort that relationship badly. The question of size of movements and their distorting effects is clearly one of empirical and statistical judgment. Since one is concerned only to isolate the effect of prices from that of all the shift parameters together, only non-offsetting changes in these latter need be considered.[37] Finally, the question of whether such changes in the other variables as are observed are autonomous or induced by accompanying price changes, is, ultimately, one of economic theory.

The question of whether the inter-war period yielded a satisfactory run of data for the estimation of either the total or the partial influence of price on the quantity of exports demanded, and, if so, which is a very complex one and cannot be finally settled in this paper. The writer's work does, however, yield some tentative suggestions in this direction :

First, it appears that the inter-war period does not offer a satisfactory opportunity for assessing the influence of the relative changes in the average values of British and competitors' exports on either total British exports or those of manufacturers, by ordinary regression techniques. This is not to imply that such relative "price" movements are without influence, but rather that they did not occur on a sufficient scale.[38]

This suggestion confirms the suspicion expressed by Marris through his hostile critic,[39] and by others.[40] It is also in line with Polak's formal findings, but in conflict with those of Chang. Neisser and Modigliani's estimates are not in direct conflict with the suggestion since they relate to only a segment of British manufacturing exports.

It must be candidly confessed, however, that the writer is not prepared to place much reliance upon these last three estimates of the influence of relative "price" movements for the simple reason that the indicators of relative "price" movements used by the several authors are of very doubtful quality.

It is clear that Polak's and Neisser and Modigliani's "price" relatives suffer from two major defects :

(1) For major countries, such as Britain, their own prices are a significant item in computing the World average ; [41]

(2) Although, in some broad sense, all exports of, say, manufactures may be substitutes, they are obviously very unequally substitutable. If, as the writer's work suggests is the case, the composition of exports of the several principal industrial countries was quite different and the prices of each country's various exports did not move together over time,[42] the "price" ratios used are not very meaningful.

The authors seem to have been aware of these difficulties.[43] They were simply concentrating on the effect of incomes rather than prices on international trade [44] and so paid relatively little attention to their price series.

Chang, who wrote earlier, was certainly aware of these difficulties.[45] He attempted to meet them both, by compiling for each exporting country : "an index of the prices of substitutes . . . obtained by taking a weighted average of the indices of export prices of a number of the countries whose exports are broadly in the same economic class as the exports from the country under consideration. The weights are given in accordance with the relative shares of the countries in total World export trade . . . in the case of the U.K. an index of prices of substitutes in the World market is compiled by taking a weighted average of the export prices, in terms of sterling, of four principal industrial competitors—U.S.A., France, Germany and Japan."[46]

If this means, as it seems to, that average values of the total exports of the four countries are considered, it would appear that Chang escaped the first difficulty above, only at the expense of greatly aggravating the second, at least so far

as the U.K. is concerned. The exports of the U.S.A. and Japan, particularly contained large amounts of raw materials which clearly were not good substitutes for British exports.

The results shown in the table are quite consistent with the criticisms just advanced. Polak, whose "price" relative for some years is, in effect, a constant, finds that he cannot estimate the effect of relative prices. The other two authors might well obtain statistically significant coefficients for relative price movements even if this writer's contention that there were not sufficient divergences of the relative average values of British and genuinely competitive exports to estimate such coefficients is correct.[47] One would expect larger divergences between two price indices if they relate to different bundles of goods.

If movements of the "prices" charged by British and competing exporters are so nearly parallel, at least for aggregates, that their effects cannot be separated, then one cannot estimate the pure partial effect of changes of British "prices" or terms of trade on British exports. One might, however, be able to estimate their effect on the assumption that relative "prices" of "bundles" of competing exports from the several sources behave as they did in the inter-war period, i.e. roughly remain constant. Such a concept of elasticity is of considerable interest, as is well known from oligopoly theory. Marris's estimate is, at best, of an elasticity of this sort. One would expect this elasticity to be lower than the pure partial elasticity estimated by Zelder and Harberger.

Another obstacle in the way of estimating the net effect of British terms of trade is the intercorrelation between those terms of trade and the income or trade of the rest of the world. While Marris's estimate of a correlation coefficient between British terms of trade and world trade of "only" about ·5 [47a] is confirmed by the writer's work, it does not follow from this that one can readily estimate the separate influence of the two. The three estimating equations below, taken from the writer's work, give some idea of the interdependence of the two estimates :

$$(1) \quad E = a_1 + \cdot 99M + e_1 \qquad R^2 = \cdot 38$$
$$\pm \cdot 37$$
$$(2) \quad E = a_2 + 1 \cdot 97P + e_2 \qquad R^2 = \cdot 94$$
$$\pm \cdot 15$$
$$(3) \quad E = a_3 + 1 \cdot 85P + \cdot 17M + e_3 \qquad R^2 = \cdot 96$$
$$\pm \cdot 18 \quad \pm \cdot 14$$

Where E are exports of U.K. produce and manufactures in millions of 1929 pounds,

M, the rest of the world's imports in millions of pounds, deflated by the same index as E,

and P is an index of British terms of trade—a ratio of the unit value index of British retained imports to that of British "home" exports—1929 = 100.

All variables are time series for the years 1924–1938, excluding 1926, and are expressed in natural logs.

It will be noted that the coefficients of both P and M are rather different when they are estimated simultaneously in (3) than when they are estimated separately in (1) and (2). This would not be the case if the two variables moved sufficiently independently of one another for their separate influences to be readily separable. Moreover, it seems that even the coefficient of P estimated in (3) contains a component more appropriately attributable to M ; i.e. that the coefficient of M can be increased to a more plausible level without appreciably reducing the degree of explanation as measured by the coefficient of determination, R^2.[48]

Whether the observed intercorrelation between British terms of trade and the rest of the world's imports [49] prevents equation (2) from yielding an "unbiased" estimate of the total as well as the partial relationship between the former and the quantity exported is a more complex question. The answer to it hinges largely upon one's judgment of the extent to which changes in the rest of the world's real income and, hence, trade were autonomous rather than induced by the accompanying changes in British terms of trade.[50]

It seems to be the generally accepted picture of the impact of the inter-war recession on world trade that the decline in aggregative demand took the form of unemployment of

resources and a drop in production in manufacturing countries and of low selling prices, i.e. adverse terms of trade, in primary producing countries.[51] Britain's customers in both sorts of economies were impoverished as a result, and her exports, whose prices were more or less maintained, declined. The same was true of manufactured exports of other countries. Their terms of trade improved at the same time because of lower import prices.

There is no need to doubt that the intercorrelation between the rest of the world's trade and British terms of trade is largely attributable to these events. It should be noted, too, that, in so far as the primary producers' volumes of imports and exports were affected by changes in the terms of trade, this connection between the terms of trade and the rest of the world's trade is precisely the "entirely valid" income effect of the foreign offer curve mentioned by Marris.[52]

The difficulty with this explanation is that it relates entirely to the downswing or, at most, (if one considers deviations from the mean rather than year to year movements) to the thirties, whereas one is really concerned with the whole inter-war period. Marris anticipated this problem by fitting his regressions for the twenties and thirties separately and finding that they yielded identical results. His results, at least, thus cannot all be explained by conditions peculiar to the latter period, even if it would have dominated the results of a regression for the inter-war years as a whole, as quite likely it would.

But the point involved in the explanation may be put in another way, more damaging to "total" estimates (and, indeed, to all estimates based on inter-war data) : The whole of the inter-war period was characterised by strong cyclical influences. One need not go all the way with Marris's "severer critics" in maintaining that any statistical relationship estimated on the basis of inter-war data reflect nothing but cyclical influences. But there is little doubt that these influences do affect such estimates. To the extent that they do, one way not be observing the relationship between, say, prices and quantities demanded, at all, in any meaningful or functional sense, but rather the concurrent response of both variables to "broad

swings in economic activity." Any such association clearly cannot be extrapolated to periods during which such "broad swings" do not occur.

Another possible limitation on the coefficient yielded by equation (2) as an estimator of the total influence of terms of trade on exports arises from the side of the balance of payments, which, as has already been suggested, may be relevant to the quantity of British exports demanded. Marris considered the possible influence of the balance of *trade* in the inter-war period and found it "not capable of serious quantitative importance."[53] A little reflection suggests, however, that once one is outside a two-goods model (i.e. once the two may diverge) it is the balance of payments, not the balance of trade which is the relevant variable, whether one wishes to consider it as a source of purchasing power additional to foreigners' production possibilities, or as a means of inducing foreigners to buy in the lending country rather than elsewhere, given the level of their imports. The writer's work suggests that the balance of payments, as measured by the official figures deflated by British export unit values, did play a significant role in explaining the level of British exports.[54] This is not in conflict with Marris's factual findings since the correlation between the British balances of trade and payments in the inter-war period is amazingly low, $r = \cdot 11$.[55] The balance of payments variable is positively correlated with the terms of trade, and if its movements were thought to be autonomous and were separated out this would again lower the elasticity of exports with respect to terms of trade.

All the preceding discussion has little bearing on the Zelder-Harberger estimates. Indeed, the writer's investigations being of an aggregative nature are of little help in assessing these estimates apart from bringing out their relation to the others. Clearly a multitude of divergences in prices of individual substitutes can be concealed by parallel movements of the average prices of British and competitive exports.

Zelder and Harberger's procedure, of course, has already been subjected to severe criticism.[56] The most telling points made were, first, that it is by no means always clear how to interpret the concept of substitution involved and, secondly,

that the transition from elasticities of substitution to those of demand is based on very shaky foundations. It is also true that neither Zelder nor Harberger seriously faced the problems of extrapolation and aggregation although both seem to be implying that their estimates are "generally" applicable.

All the preceding discussion dealt with whether or not it is possible to estimate total or partial relations. Before leaving this matter it is desirable to turn briefly to the logically prior, but practically less interesting, question of which of the two estimates one would prefer to have if one had the choice. No doubt, the perfectionist answer is that one would prefer to know the partial relations between the quantity exported and each of its determinants as well as all the partial relationships *among* the explanatory variables.

Short of this degree of perfection, the answer clearly depends upon the problem with which one wants to deal. For problems of commercial policy one wants price elasticities appropriate to (assumed) competitors' reaction. Thus, when Marris is striving for the "pure" partial elasticity he assumes, by implication, that Britain's competitors would not follow her lead in improving their terms of trade by protectionist devices. Similarly, the use to which Zelder and Harberger put their elasticity implies the doubtful assumption that devaluations would not be followed. The Marris elasticity, calculated on the assumption that competitors "follow the lead," may have been conceptually more appropriate for their purpose.

In principle, there is, of course, no reason why one should stop at incorporating oligopolistic price reactions in the elasticity used. The whole nexus of induced reactions may be so incorporated by using the total elasticity. The principal objection to such a use of the total concept is doubt as to the stability of the induced relationships : It is fairly easy to persuade oneself that price or income elasticities of demand for exports are reasonably stable over time. It is much more difficult to believe that incomes or tariffs abroad respond in a regular manner to changes in British terms of trade, or even that competitors' terms of trade do so.[57]

This brings one again to the final problem to be considered in this paper : Are these inter-war estimates of any relevance

to present conditions? Again, no definite answer can be given. The general problems of statistical inference from time series are too well known to require discussion here, and too complex to permit it. It need only be mentioned that the "hidden premise of the uniformity of nature" is perhaps more than usually strained in the present case. Clearly the economic situation, and perhaps most of all Britain's international economic situation, is quite different now from what it was in between the wars. It requires either very considerable faith or some definite demonstration to convince one that the results are still applicable. Marris, of course, attempts the demonstration, but not all will find it convincing.

All in all, and perhaps this is the real moral to be drawn from this review, it is a fortunate thing that the day is now not too distant when the accumulation of reliable post-war data permits workers in this field to face their very formidable estimating problems without the additional complication of attempting, initially, to find something they do not really seek.

S. F. Kaliski

Carleton University, Ottawa.

REFERENCES

1. Work on this paper was started while the writer was Research Fellow in Economic Statistics at the University of Manchester. It formed part of a larger project largely carried out at the Department of Allied Economics, Cambridge. Some of the results of this project were reported in "Extraneous Estimates and Goodness of Fit: A Problem in Empirical Economics," *Manchester School*, vol. XXVIII, No. 1, (January, 1960), pp. 59-71. The writer has been greatly helped by Professor J. Johnston's comments on an earlier draft.
2. For a recent survey of the literature see Hang Sheng Cheng, "Statistical Estimates of Elasticities and Propensities in International Trade: A Survey of Published Studies," *I.M.F. Staff Papers*, Vol. VII, No. 1, (Apr., 1959), pp. 107-158; and "A Collection of Statistical Estimates of Elasticities and Propensities in International Trade," mimeo., I.M.F., Sept., 1957.
3. T. C. Chang, *Cyclical Movements in the Balance of Payments*, Cambridge, 1951.
4. H. Neisser and F. Modigliani, *National Incomes and International Trade*, Urbana, 1953.
5. J. J. Polak, *An International Economic System*, London, 1954.

6. Eg. by A. J. Brown, "The Rate of Exchange," in T. Wilson and P. W. S. Andrews, (eds.), *Oxford Studies in the Price Mechanism*, Oxford, 1951 ; F. Machlup, "Elasticity Pessimism in International Trade," *Economia Internazionale*, Vol. III (1950), pp. 113-137 ; D. J. Morgan and W. J. Corlett, "The Influence of Price in International Trade, A Study in Method," *Journal of the Royal Statistical Society*, Series A, Vol. CXIV, Pt. III, (1951), pp. 307-352 ; and G. H. Orcutt, "Measurement of Price Elasticities in International Trade," *Review of Economists and Statistics*, Vol. XXXII, (May, 1950), pp. 117-132.
7. See, eg. R. Stone, *The Measurement of Consumers' Expenditure and Behaviour in the United Kingdom*, 1920-1938, Vol. I, Cambridge, 1954, Chapt. XIX ; H. Theil, *Economic Forecast and Policy*, Amsterdam, 1958, Chapt. VI ; H. Wold, *Demand Analysis*, New York, 1953, Chapt. 2 ; D. Cochrane and H. G. Orcutt, "Application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms," *Journal of the American Statistical Association*, Vol. 44, (Mar., 1949), pp. 19-61, esp. p. 57, on various aspects of this matter ; see Kaliski, *loc. cit.*, esp. pp. 59-60, 62-63 for a summary of some of the views.
8. *Loc. cit.*, p. 90.
9. R. L. Marris, "The Purchasing Power of British Exports," *Economica*, N.S., Vol. XXII, No. 85, (Feb., 1955), pp. 13-28, and "Rejoinder," *ibid.*, Vol. XXIII, (Feb., 1956), pp. 67-70.
10. *Op. cit.*
11. R. E. Zelder, "Estimates of Elasticities of Demand for Exports of the United Kingdom and the United States, 1921-1938," *Manchester School*, Vol. XXVI, No. 1, (Jan., 1958), pp. 37-47, and "Elasticity of Demand for Exports, 1921-1938," unpublished Ph.D. dissertation, University of Chicago, 1955.
12. A. C. Harberger, "Some Evidence on the International Price Mechanism," *Journal of Political Economy*, Vol. LXV, No. 6, (Dec., 1957), pp. 506-521 ; and *Review of Economics and Statistics*, Vol. XL, No. 1, Pt. 2, (Supplement, Feb., 1958), pp. 123-132.
13. But too much play should not be made with this distinction: Explicitly, at any rate, Mr. Marris is considering the problem of the optimum degree of restriction and Dr. Zelder and Professor Harberger that of exchange stability. Clearly, one can be optimistic in *both* these contents. See, however, H. G. Johnson, *International Trade and Economic Growth*, London, 1959, Chapts. VI (esp. pp. 162-8), and VII (esp. Secs. III and IV), and "The Balance of Payments," *Pakistan Economic Journal*, Vol. VIII, No. 2, (June, 1958), pp. 16-28, esp. pp. 25-28 for a comparison of devaluation and trade restrictions as policies for correcting a balance of payments deficit.
14. The point that marginal costs of trade are *increasing* and this ought to be borne in mind in formulating policy was made earlier by E. A. G. Robinson and Marris, "The Use of Home Resources to Save Imports," *Economic Journal*, Vol. LX, No. 237 (Mar., 1950), pp. 177-181.
15. Imports plus exports.
16. Polak's equations appear to be linear in the indices. The coefficients are then identified as mean elasticities for the period, presumably by taking these indices to an appropriate base. *Cf.* Cheng, mimeograph, p. 165.
17. *Op. cit.*, p. 160.

17a. That is, which differ from zero by a larger amount than is likely to result from mere sampling fluctuations.

18. *Op. cit.*, p. 72.

19. *Op. cit.*, p. 51.

20. *Op. cit.*, p. 96.

21. *Op. cit.*, Table 1.

22. See below for an explanation of these terms.

23. Italy, Japan, Sweden, Belgium, Switzerland, Czechoslovakia, and Austria.

24. For the primaries countries the purchasing power of exports and of industrial production, and capital imports are used in place of income.

25. See below, p. 10.

26. Morgan and Corlett, *loc. cit.*, p. 318, note.

27. Harberger, *Review of Economics and Statistics*, pp. 123-124. Some of Harberger's critics agree with this judgment, e.g., H. Neisser, "Comment," *ibid.*, pp. 129-132, esp. 130.

28. *Loc. cit.*, Table 3.

29. Zelder, *loc. cit.*, p. 44.

30. Harberger, *Review*, p. 125.

31. The mathematically inclined reader will have no difficulty in following through this analogy to the total differential and partial derivative, respectively, of the function expressing the quantity exported in terms of its determinants.

32. See M. F. G. Scott, "Interdependence and Foreign Trade," *Oxford Economic Papers*, Vol. 19, No. 1, (Feb., 1957), pp. 88-107, and Neisser and Modligliani, *op. cit.*, pp. 19-22, 75-91, 97-101, for similar distinctions between partial and total influences. Scott points out, quite correctly, that, although an unambiguous meaning can be attached to the total relationship, there are as many partial relationships as one cares to consider. Only the simplest of these is considered above ; see however below, p. 16.

33. Unless one chooses the Zelder-Harberger alternative.

34. For a fuller exposition of the statistical theory see e.g., Stone, *op. cit.*, pp. 302-303.

35. *Cf.* Marris, Article, p. 27.

36. In non-trivial cases. The two trivial ones—when the movements of the several explanatory variables are uncorrelated and when variables other than price do not change have already been mentioned.

37. *Cf.* Marris, Article, p. 20, note. In the present case non-offsetting may also be interpreted in an additional, and more subtle, sense : Since the total quantity of British exports is necessarily in principle an index number, changes in the shift parameters which affect only the composition or destination, but not the "total," of this quantity may be neglected.

38. There were some quite large divergences in the movements of British and competitors' export "prices" in the 1930's. But these all occurred at the times of the major currency devaluations of the period and seem to have been accompanied by peculiar, and offsetting, developments in commercial policy.

39. *Article*, p. 21.

40. Eg., Harberger, *Review*, p. 123.

41. In Polak's case this difficulty is quite grotesquely magnified. In order to push his data back from 1929 (when the published League of Nations series on volume and value of trade are first subdivided) to 1920 he relied on data supplied him by the late Mr. F. Hilgerdt of the League (Polak, p. 174). These data were kindly made available to the writer by Mr. W. R. Leonard of the United Nations Statistical Office. Mr. Leonard warned, and the writer confirmed, that the deflator applied by Mr. Hilgerdt to his trade in manufactures prior to 1929 was Schlote's index of British export prices. The non-manufactures were deflated by the Sauerbeck index of British wholesale prices. (W. Schlote, *British Overseas Trade from* 1700 *to the* 1930's, Table 26). Now, however justifiable this deflating procedure may have been for its original purpose of providing estimates of volume of trade by major commodity groups, it is not surprising that Polak did not get satisfactory results for the influence of relative prices on British exports of manufactures by using the deflator as denominator. On the other hand, Neisser and Modigliani seem to have constructed their series with great care. (*Op. cit.*, Technical note to Part 3).
42. On this last question see R. E. Zelder, *loc. cit.*, pp. 39-40, and Kaliski, "A comment on Dr. Zelder's Estimates," *Manchester School*, Vol. XXVI, No. 3, (Sept., 1958), pp. 241-246, esp. pp. 241-245.
43. See Polak, *op. cit.*, pp. 51, 65; Neisser and Modigliani, *op. cit.*, pp. 47, 51, 300.
44. Polak, *op. cit.*, p. 59, Neisser and Modigliani, *op. cit.*, pp. 1, 5, 6, etc.
45. Chang, *op. cit.*, pp. 48-49.
46. *Op. cit.*, p. 49. The matter is not put any more clearly in Chang's unpublished thesis of the same title. But the language and what is known of the availability of data together strongly support the interpretation which follows.
47. This contention is based upon the examination of an index of competitors' prices prepared for the purpose. For 1924-32 this index is based upon unit value indices for some 90 groups of the German trade classification and supplementary Japanese indices for cotton yarns, cotton tissues, and silk and artificial silk tissues. No adjustment is made for the fact that the individual commodities comprising each of the German-Japanese groups are combined in different proportions than in corresponding aggregates from the British trade classification, but the several groups are combined with British current weights. (I.e. the German-Japanese groups are, in effect, treated as homogeneous commodities and problems of weighting within them are ignored). For 1933-38 this index is extrapolated by the unit value index of U.S. exports of Finished Manufactures. The resulting index may be said to be less comprehensive, but more nearly comparable to the British than those used by Polak, Chang, and Neisser and Modigliani.

47a. Rejoinder, p. 69. This means, roughly speaking, that only 25 per cent. of the variation in each of these variables can be expended in terms of corresponding movements of the other variable.

48. Kaliski, *Extraneous Estimates*.
49. One gets much the same results by substituting the rest of the world's exports for its imports.
50. *Cf.* e.g. Marris, Article, p. 20.
51. *Cf.* e.g. Marris, Article, p. 20; D. J. Morgan and F. W. Paish, "The Purchasing Power of British Exports Further Considered," *Economica*, N.S., Vol. XXII, (1955), pp. 329-335, esp. p. 330; Neisser and

Modigliani share the view that income changes spread from industrial to "primaries" countries. But they view the inter-war period as a whole, and, therefore, are less insistent about the purchasing power of the latter countries being affected by the changes in the prices rather than the volumes of their exports (*op. cit.*, p. 5). Their figures (*op. cit.*, Tables 36, 38, T4) do suggest however that during the recession price changes were more important than changes in volume. The prices of food and raw materials imported by industrial from "primaries" countries declined by about 40 and 50 per cent., respectively between 1929 and 1932. The volume of the food imports did not decline at all during this period, but that of raw materials imports declined by about 20 per cent. On the other hand, between 1925 and 1929 changes in price were more than compensated by those in volume. The prices of food and raw material imports fell by some 8 and 20 per cent., respectively. But the volumes of these exports increased by some 25 and 30 per cent., respectively.

52. Article, p. 20.

53. Article, p. 20.

54. The numerical results, for what they are worth, are that an additional pound loaned abroad increased exports by about seven or eight shillings at given levels of the rest of the world's imports (exports) and prices. The partial regression coefficient is ·36-·39·07, depending upon exactly what version of the other variables is included simultaneously.

55. *Cf.* the suggestion of Neisser and Modigliani, *op. cit.*, p. 25, that the balance on invisible items "has generally fluctuated within narrow limits" and so did not affect changes in the balance of payments. This does not seem to have been the U.K. experience as measured by the official statistics. But it must be admitted at once that these statistics are of somewhat doubtful quality. See e.g. A. E. Kahn, *Great Britain in a World Economy*, London, 1946, p. 296, and T. E. Gregory, "Foreign Investment and the British Balance of Payments," in *Foreign Investment, Lectures on the Harris Foundation*, Chicago, 1928, pp. 120-125.

56. Especially by Neisser *loc. cit.*, and W. L. Smith in the same discussion, pp. 127-129. The basic procedure was earlier subjected to detailed criticism by Irving Morrisset, "Some Recent Uses of Elasticity of Substitution, A Survey," *Econometrica*, Vol. 21, (Jan., 1953), pp. 41-62, but Zelder, at least, seemed confident that he could escape Morrissett's strictures (*loc. cit.*, p. 37). Both Zelder's basic estimates and his transition procedure have also been specifically criticised by the writer in the *Manchester School* comment already mentioned.

57. See above re protection and devaluation. This, incidentally, is an interesting aspect of the Marris-Morgan and Paish (Marris, Article and Rejoinder, Morgan and Paish *loc. cit.*) controversy which was neglected by both sides. It may be that it *does* matter whether British terms of trade are altered from the side of import prices or that of export prices because competitors will react differently to the two sorts of alteration. They may, for instance, accept an improvement in their terms of trade from the side of import prices, as they did during the thirties, but decline to follow increases in British export prices. Marris's calculated elasticity is more appropriate to the former than to the latter behaviour, and this is of some importance for appraising his policy recommendations if one believes that British commercial policy is to-day more likely to have an impact on the prices of exports than on those of imports.

Public Expenditure and Revenue: An International Comparison*

1. Introduction

The research and analysis which follows in this paper owes much of its inspiration to a fairly recent article published by A. M. Martin and W. A. Lewis in *The Manchester School.*[1] The authors of that interesting paper began their study with the intention of examining the relationship between income (or gross national product per capita) and the levels and patterns of public expenditure and revenue. But Martin and Lewis never quite fulfill that promise made to the reader for in the analysis only classification and casual empiricism result. The object of this study is to attempt a more rigorous statistical test of the relationship which they suggest in their article and some others which they seem to have relegated to the less important.

More precisely, the concern of this article is to determine the relationship, if any, between the levels of per capita income and the share of public expenditure (or revenue) in gross national product and the sources of this revenue : that is to say, are the levels of government expenditure and the sources of revenue primarily policy determined by domestic social and political history or is there some economic law which reveals a relation between public expenditure and income levels much like Engel's law of consumption expenditure.

We have already mentioned one criticism of the Martin and Lewis attempt ; that there is no use of regression, or extensive rank correlation, analysis in that paper. The method here employed, however, is an application of simple univariate regression functions to explicitly determine acceptance or rejection of the hypothesis. Secondly, it has also been suggested that the Martin and Lewis results are heavily biased towards known

*This paper was originally published, with minor changes, as Memorandum B-9 under the *Stanford Project for Quantitative Research in Economic Development* in July, 1960. The author wishes to express his thanks for the help given by Professors Houthakker and Chenery. They are, of course, in no way responsible for the errors in method or implication which may still remain.

[1]A. M. Martin and W. A. Lewis, " Patterns of Public Revenue and Expenditure," *The Manchester School of Economic and Social Studies*, September 1956, pp. 203-244.

welfare nations and as a result their examination does not reveal a clear relationship between income levels and public revenue and expenditure which may indeed exist. Their sample size is a small one (N=16) while the sample used in this study is more than 30 in almost all cases. But nine out of sixteen of their sample are either members of the British Commonwealth and past British Colonies, or are present British Colonies ; six are high income nations representing Western Europe ; and the remaining observation is a lone representative of Latin America. One might suppose, a priori, that this sample would bias the results towards welfare nations and thus might hide support for a relationship between public expenditure and per capita income, and for sources of government revenue as well.

Another serious objection that I have to the Martin and Lewis sample is their method of choosing the country observations for the international cross-sectional comparison. In all cases but one, Columbia, the sample was chosen for *one year*—1953 or 1954. This seems to the present writer an unfortunate method of sample choice since they (and I) are primarily concerned with comparisons at different levels of economic development and by choosing one year we may introduce the effects of the business cycle upon the long run relationship and obscure meaningful results. If we choose 1953, for instance, we would find a smaller share of current government expenditure in national income since it is a boom period in most countries (and proportionately more so for rapidly developing nations, those which seem to exhibit more violent cyclical swings, as well as those which are actively pursuing anticyclical policies ; all of which suggests a bias against the wealthy). We might also expect that choosing 1953 in some cases (a peak in United States activity and elsewhere) and 1954 in others (a recession year in the United States and in many other countries) might bias our results although whether this would strengthen or weaken the statistical fit is not, a priori, clear. In order to avoid this obvious problem, I have eliminated the business cycle as best possible (given the constraint of the data) by summing the observations[1]

[1]These data are taken from the useful compilations in the *Yearbook of National Accounts Statistics*, 1958, Statistical Office of the United Nations, Department of Economic and Social Affairs, New York, 1959.

for each country over the 1950's ;[1] in almost all cases the span is from 1951 to 1957.

By taking a larger sample, and one which I hope has helped eliminate short run cyclic effects of the business cycle, I have made a sacrifice. It was not possible for me to give a breakdown of the directions of government expenditure which the Martin and Lewis article achieves. While improving the sample size, we must be content for the present with an examination of such aggregates as total government revenue, current government consumption, and total government revenue less defence expenditure. Fortunately we do not have to make the same sacrifice with respect to public revenue sources for this present analysis, to my mind, attains a higher degree of disaggregation than that of the Martin and Lewis attempt.

2. Public Expenditure and Revenue

A priori, does economic theory suggest a positive correlation between the share of government activity (in both revenue and expenditure measurements) and the level of per capita income? Why should we conceive of government expenditure as a luxury good revealing an income elasticity greater than unity? In negation of such an hypothesis, we might appeal to the post-war growth models presented by such " balanced growth " economists as Nurkse and Rosenstein-Rodan,[2] as well as those economists who are preoccupied with the importance of social overhead

[1] I should frankly admit at the outset that short-run fluctuations have not been entirely eliminated from the test sample. Whenever we deal with percentage shares—ratios of government expenditure or revenue to national income—both the numerator and the denominator, which are given in domestic currency, have been averaged over as many years in the 1950's as available. However, in the regression tests themselves, the income variable must be in comparable units of measurement. I therefore took advantage of a recent study which gives a cardinal ranking of per capita incomes in terms of United States dollars (see footnote 1, p. 48). These figures, unfortunately, are for 1953 alone and so may reflect short run movements. Although the percentage shares computed and used throughout this study are unsullied by cyclic distortion, the income variable may not necessarily be unaffected by it. It does seem, however, that the cardinal rankings of per capita income would be little affected due to the extremely large spread of income levels in the cross section.

[2] R. Nurkse, *Problems of Capital Formation of Underdeveloped Countries*, Basil Blackwell & Mott, Ltd., Oxford, 1953, pp.58, 63-67, etc.
P. N. Rosenstein-Rodan, " Problems of Industrialization of Eastern and Southeastern Europe," *Economic Journal*, 1943, pp. 202-211.

and the social investment base. It has been suggested by this school that marginal capital output ratios are extremely high among low income nations due to the complete lack of adequate electrification, power, transportation and urban housing bases on which to build a private prosperity. Thus, in capital goods expenditure at least, this would imply high government expenditure shares in low income and developing nations relative to the middle-wealthy income groups who have already made that " lumpy " investment and whose marginal public capital outlay concomitant with output growth is much less. This is the only really negative suggestion I could find in the literature with respect to the hypothesis.

In informal discussions of the Martin and Lewis article, there has been some reference to the stagnation thesis ; that it would support the existence of an elasticity of government revenue-expenditure, with respect to income, greater than unity. The argument goes that with the historical process of development, there is a shrinking of investment opportunities and in order to support a Keynsian full employment level of income, an increasing amount of government expenditure is necessary to compensate for the dearth in private investment outlays. This hardly seems, to the present writer, an adequate argument in support of the thesis that the *share* of government expenditure increases over time with rising income levels : certainly the absolute level of government expenditure will increase but whether the share increases is not obvious. We must, it seems to me, appeal to other arguments to give support for a positive correlation.

One reason for suspecting an increasing government share results from the Kuznets-Clark studies on sectoral productivity. It is generally agreed that the government sector is characterized by slow growth in productivity. Government expenditure is in the form of public services and both the Kuznets and Clark studies[1] suggest that the rate of productivity increase in the service industries is certainly less than manufactures and even

[1]See for instance Colin Clark's *The Conditions of Economic Progress*, Macmillan & Co., Ltd., London, 1951.
And H. B. Chenery, " Patterns of Industrial Growth," Stanford Project for Quantitative Research in Economic Development, Stanford, 1959 (mimeo).

less than that of agriculture. Then if public services are a function of output, a greater share of government expenditure would result with increases in income or average labor productivity. It is also true that labor scarcity increases as income levels rise with the process of capital accumulation; and thus, labor becomes more expensive over the development process. But the government sector is also characteristically extremely labor intensive. This would also give strong support to an hypothesis of positive correlation.

One of the characteristics of economic development and the course of industrialization is the urban movement and with it disappears the informal security of the family and village. Urban populations must be supplied with formal social security, unemployment insurance, a complex cosmopolitan political machine and formal community protection to replace the family and village functions. Then, again, this would support the hypothesis of an increasing share of government expenditures in national income.

Finally, during the process of economic development there should be similar development in progressive social values. In most of the development discussions in the post-war literature, for instance, there has been implied a correlation between social and political progressiveness and income levels. "Progressiveness" is one of the parameter changes in the Rostow system, as interpreted by others, that is coincident with the take off; and most of the endless discussions of equilibrium traps and vicious circles apply to social and political progressiveness, or the lack of it.[1] This suggests a correlation between degrees of welfarism and levels of income. This argument, it seems to the present writer, might be more applicable to a 19th century system for it seems that many underdeveloped nations are yielding to the demonstration effect in adapting Russian or modern European planning experience to their own development problems.

3. The Public Share: The Tests

In the tests themselves, three types of measurement were used to determine the government share in income;[2] current

[1] For an extreme view, see Paul A. Baran, *The Political Economy of Growth*, Monthly Review Press, New York, 1957, pp. 163-249.

[2] In this paper, government refers to an aggregation of all government levels including municipal, state, federal, etc.

expenditure on public services (G_e), total current revenue (R), and total current revenue less the expenditure on national defense (R_d). Both G_e and R were tested separately simply in an attempt to isolate the role of the public sector as a saver in the system. In all cases, a double logarithmic function was used of the form

$$\log R/Y = \log a_o + a_1 \log Y_c$$

or

$$R/Y = a_o \; Y_c^{a_1},$$

where Y is gross national product in the ith country's current prices, R is some measure (of the three) of government expenditure or revenue in current prices, and Y_c is per capita income in United States 1953 dollars.[1] This log function does give a better fit in all cases relative to a linear function. It also, of course, has an important economic interpretation since a_1 is a measure of constant elasticity.[2]

The best regression in terms of fit and correlation coefficients results ($N=33$; $r=\cdot 73$) when we measure the government share by use of total revenue divided by gross national product (R/Y) ;

$$R/Y = \underset{(\cdot 048)}{\cdot 0378 \; Yc^{\cdot 290}}$$

When we use current government expenditure (G_e/Y) in determining the public share in gross national product we get a little poorer correlation but the difference is small ($N=33$; $r=\cdot 70$).

[1]These per capita rankings are for 1953 and were taken from C. P. Kindleberger, *Economic Development*, The MacGraw-Hill Book Company, Inc., New York, 1958, Table 1.1, p.6.

[2]It would have been interesting also to test a hyperbolic function but time has proved too much a constraint. A hyperbolic function of a type that approaches an asymptote parallel to the per capita axis, would test the Colin Clark hypothesis that there is some taxable capacity limitation—his suggestion was 25% which has been negated in this data sampling at least since Austria is over 30% (see Figure 1). Here, however, we chose a relatively simple function which seemed to fit better than a linear function since we are primarily concerned with the evidence of a positive correlation between per capita income and the government share, but the double log junction gives us a measure of elasticity as well. Clearly, as a predictive device, the double log function presents some problems (problems which are outside the scope of this sample even so) since the government share can exceed 100% at high levels of per capita income : there should be some cut off point.

Another interesting function to test might have been an inverse trigonometric function which has an inflexion point and approaches two asymptotes inversely (Figure 2). This would suggest a minimum subsistence level of government expenditure share in national income.

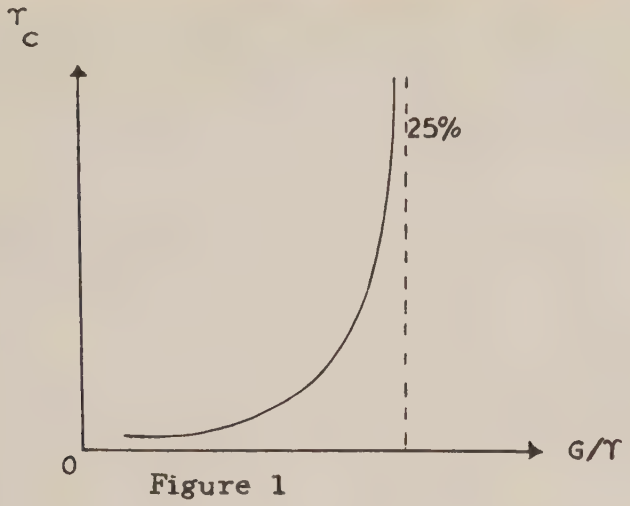


Figure 1

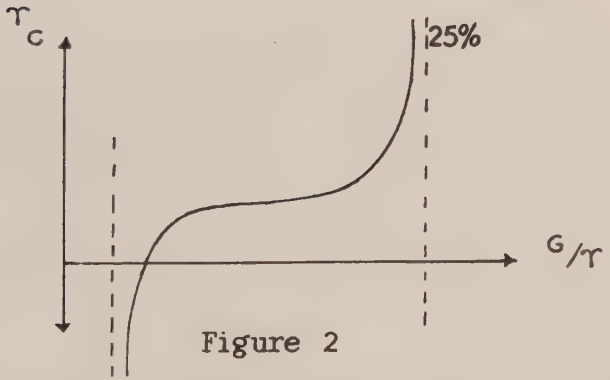


Figure 2

In both cases (the standard errors are small) the correlation coefficients are significant for $P = \cdot 99$. Interestingly enough the estimates for the exponent parameters are extremely close for both fits; for R/Y and G_e/Y, α_1 is respectively ·290 and ·257. Thus when we use current expenditure and total revenue as measures of the public share in gross national product, there is a definite positive correlation between per capita income and the government *share*. From low income to middle income the share increases rapidly with respect to per capita income while the rate

of change lessens with movements to high income nations : the income elasticity of government expenditure and revenue is greater than one, 1·257 and 1·290 respectively. These results are far brighter than those of the Martin and Lewis sample which produced a coefficient of rank correlation of only ·46.[1]

TABLE 1

	Correlation coefficient	Sample size
(1) $R/Y = \cdot 0378 Y_c^{\cdot 290}_{(\cdot 048)}$	$r = \cdot 73$	$N = 33$
(2) $G_e/Y = \cdot 041 Y_c^{\cdot 257}_{(\cdot 047)}$	$r = \cdot 70$	$N = 33$
(3) $R_d/Y = \cdot 0357 Y_c^{\cdot 272}_{(\cdot 092)}$	$r = \cdot 53$	$N = 29$

Incidentally, a test was made on the relationship between the share of current public saving in total revenue, as well as the share of public saving in gross national product (where public saving is defined in the usual way, current revenue minus current expenditure), and levels of per capita income. The suggestion here is that the underdeveloped nations need public financed saving to offset the relatively small private saving propensity thus precipitating an inverse relationship with per capita income ; or that degrees of socialism (new capital created in the society without a corresponding increase of private wealth[2]) vary with income levels within the cross-section. In both tests independence between the two variables was exhibited. But unfortunately these figures for " degrees of socialism " do not include those profits which are ploughed immediately back into public owned industries and firms. Thus our results may not be entirely meaningful.

Although both G_e/Y and R/Y reveal a positive correlation with per capita income, R_d/Y does not exhibit as good a fit. That is, when we eliminate defense expenditure from total revenue the correlation coefficient between R_d/Y and Y_c is

[1]Martin and Lewis, p.206.

[2]Martin and Lewis, p.216.

significant at the 99% level ($N=24$, $r=\cdot53$), but the least squares fit is significantly reduced ($a_1=\cdot272$ and thus the elasticity of R_d with respect to income is 1·272). The general result seems to be that when defense expenditures are eliminated from R, the high income countries (which are also the high defense expenditure nations in terms of share) find their public share falling relative to the smaller nations. This would not be an unpalatable conclusion if we assume that society votes on an equilibrium solution between *aggregate bundles* of public and private goods and that defense expenditure is determined externally. The only other conclusion is that " community " utility functions are shaped in such a way that the elasticity of defence expenditure with respect to public expenditure is greater than unity ; i.e., the higher the income the greater the propensity toward military imperialism and/or the concomitant fear of aggression.[1]

4. Sources of Public Revenue

In the first section of this study we found that there is indeed a positive correlation between levels of per capita income and the public share in gross national product. But an equally important question to ask might be : Is there any systematic relation between the methods of obtaining tax revenue and the levels of income, or are the sources of tax revenue dictated mainly by policy decisions independent of income levels ? This portion of this paper disaggregates total government revenue into three sources ; corporation profits taxes ; direct taxes on households which include income taxes, poll taxes, social insurance taxes, land taxes and death duties ; and indirect taxes such as excise taxes, foreign trade duties and sales taxes. Martin and Lewis separate total revenue into direct taxes and foreign trade duties leaving the rest as a catchall residual. This hardly seems very helpful especially since one would be certain to find the size of foreign trade greatly influencing that source of government revenue. One further point of dissimilarity between

[1]It may be of interest to note that there are four nations which are relatively rather far off the regression line ; United States, Canada, and Australia have low percentage shares relative to per capita income while Austria reveals high percentage shares. This seems to hold true for both the test of G_e/Y and R/Y.

this study and that of Martin and Lewis is that we measure the importance of the revenue source by the share of each component in the figure for total current revenue. The study by the authors above measures this share by the ratio between each component and *national income.* Since they find poor results in the correlation between R/Y and Y_c then we would certainly expect the same results or worse by using national income as the divisor in our revenue source measures. Nor does this measure really isolate the importance of each component as a source of government revenue. In order to eliminate this dependence between the two tests, I have used total government revenue rather than national income as a base by which to measure the relative importance of these three revenue sources.

As before, in all these tests we have used the double logarithmic function and in the initial regressions the sample size varies from 30 to 32 countries.

TABLE 2

	Correlation coefficient	Sample size
(1) $C_t/R = \cdot039Y_c^{\underset{(\cdot089)}{\cdot210}}$	$r=\cdot41$	$N=30$
(2) $D_t/R = \cdot034Y_c^{\underset{(\cdot089)}{\cdot366}}$	$r=\cdot61$	$N=30$
(3) $I_t/R = 1\cdot483Y_c^{\underset{(\cdot041)}{-\cdot183}}$	$r=\cdot63$	$N=32$

The exponent of the C_t function does indeed have a positive value ($a_1=\cdot210$) but the fit is poor. Using a sample size of thirty, however, the correlation coefficient is significant at the 98% level ($r=\cdot41$).[1]

[1]In the test on corporation profits taxes, there are seven nations which are relatively far off the regression estimate; Denmark, Belgium and France have extremely low shares relative to the logarithmic fit and Union of South Africa, Peru, Portugal, and Japan exceed on the extreme high side. Perhaps this may suggest a second variable which could be used with income per capita or with it in a multivariate regression, but it is not immediately clear to the present writer. This is discussed more generally in the text that follows.

Using this same sample ($N=30$), I applied the test on direct taxes share (D_t/R) with much better results. We find a high positive correlation between D_t/Y and Y_c with $a_1=\cdot366$ and $r=\cdot61$ which is indeed significant at the 99% level. Income taxes, death and asset duties, and land taxes are by far the greater part of D_t (excluding foreign trade taxes). As far as death and asset duties are concerned, their yield may be low for low income countries since the ratio of capital to income is thought to be low,[1] but the evidence of a positive correlation between income level and K/O is far from clear. Land taxes may also produce low revenue for low income nations since large landowners control political decision making (i.e., India and Latin America). The income tax makes up a large share of D_t. Thus the high correlation between D_t/Y and Y_c suggests in turn a high correlation between the levels of income and the progressive nature of the tax. (Or it may simply reveal either that (1) income taxes are more difficult to collect than indirect and that there are economies of scale in the collection of direct taxes or (2) that low income countries find it difficult to collect an adequate revenue from the income tax with per capita incomes so low and close to subsistence or (3) the unequal distribution, and small percentage of high income receivers is not subject to tax since they themselves are the decision making body.) In any event, it is clear that there is a positive correlation between the share of direct taxes in total government revenue and the level of Y_c.

That the correlation between income levels and the direct tax share should be better than the results using profits tax share (in terms of goodness of fit) should not be too surprising. Even if we take the tax rate on profits as given and equal at all income levels, there is no reason to suppose that share increases with income levels. On the contrary, historical evidence, in spite of the Marx-Kalecki models, supports an increasing labor share and a falling profit share. Even though the incorporated share of industry increases with income, then, an increasing profit share does not seem to be a likely source of a positive correlation between income levels and the share of C_t in government revenue. Rather, the *rate* of tax seems to be the most likely cause of our results.

[1]Martin and Lewis, p.224.

The best fit of the three results from the regression of I_t/R on Y_c. Using a sample size of 32, we find a negative correlation between the levels of per capita income and the relative importance of indirect taxes in total government revenue ($\alpha_1 = -\cdot 183$; $r = \cdot 63$). Thus it seems that with the process of economic development and rising income levels there are relative shifts out of indirect taxation and into direct taxation (which, in this latter sense, includes both levies on households and corporations). Not only may this reflect an increasing degree of progressiveness, but it may also indicate the political instability of the underdeveloped nation and the necessity of imposing inconspicuous tax duties, as well as reflecting the relative ease with which indirect taxes may be collected in low income nations.[1].

JEFFREY G. WILLIAMSON

[1] In the tests on the indirect tax share, Canada, Australia, the United Kingdom and the United States are all extremely high for their income levels while Peru and Burma are quite low. The same culprits arise in the direct taxes test. Canada, Australia, the United States and the United Kingdom all reveal low shares (relative to income levels) while India, South Korea and Austria yield high shares.

TABLE 3

Country	Span	C_t/R^*	I_t/R^*	D_t/R^*	R/Y	G_e/Y	R_d/Y	Y_c
Australia	51–57	·16	·49	·35	·22	·16	·18	921
Austria	51–57	·07	·41	·52	·33	·27	·31	290
Belgium	51–57	·09	·42	·49	·21	·22	·17	717
Burma	51–57	—	·47	—	·18	·14	—	43
Canada	51–57	·21	·52	·27	·24	·20	·17	1,318
China	51–57	·09	·83	·08	·19	·18	—	160
Columbia	51–57	·13	·73	·14	·14	·14	—	250
Costa Rica	51–57	·15	·73	·12	·18	·11	—	180
Denmark	51–55	·06	·46	·48	·23	·20	·19	740
Ecuador	51–56	·08	·76	·16	·19	·15	·16	100
Finland	54–57	·13	·52	·35	·20	·22	—	450
France	51–57	·07	·55	·38	·30	·28	—	600
W. Germany	51–57	·09	·41	·50	·35	·27	·31	482
Greece	51–57	·10	·61	·29	·20	·20	·14	174
India	50–55	·05	·67	·28	·09	·07	·09	60
Ireland	51–57	·10	·70	·20	·21	·22	·20	416
Jamaica	51–56	·14	·75	·11	·11	·10	·11	170
Japan	51–57	·17	·48	·35	·21	·14	·21	197
S. Korea	55–57	·04	·61	·35	·09	·12	·03	70
Luxembourg	52–57	·20	·31	·49	·30	·24	·28	700
Netherlands	52–57	·14	·39	·47	·29	·23	·23	600

*In order to insure comparability and preserve additivity, the R's are the sum $C_t+I_t+D_t$ and not strictly total government revenue.

TABLE 3—*continued*

Country	Span	C_t/R^*	I_t/R^*	D_t/R^*	R/Y	G_e/Y	R_d/Y	Y_c
New Zealand	51–57	·19	·32	·49	·27	·22	·24	968
Nigeria	52–56	·11	·71	·18	·09	·06	·09	62
Norway	51–57	—	·46	—	·29	·22	·25	717
Panama	53–56	·14	·64	·22	·27	·23	·27	301
Peru	51–56	·35	·46	·19	·12	·11	—	140
Phillipines	51–57	·09	·82	·08	·10	·09	—	90
Portugal	51–57	·22	·48	·30	·17	·14	·14	250
Sweden	51–57	·14	·32	·54	·30	·24	·25	910
Union of S.A.	53–57	·31	·42	·27	·18	·19	—	283
United Kingdom ...	51–57	·17	·49	·37	·26	·26	·18	930
United States	51–57	·21	·33	·46	·24	·22	·13	1,908

Administrative Overheads in British Manufacturing

Introduction.

There seems no reason to believe that the labour market is immune from the shocks of the process of "creative destruction," or that this process is unaffected by the size and structure of the labour force. New products, techniques, and wealth, must involve new skills, trades, and conditions of employment, as well as the more familiar movement between existing occupations, industries, and their locations. The speed of the working population's adjustment to these changing demands might hasten or slow-down the rate of innovation, or determine its character and direction. Yet apart from a few generalisations about the importance of changes in the level of real wages in models of economic growth, little attention is paid to the dynamic characteristics of the labour force. This is partly, perhaps, because so little is known, or can be discovered in detail about the impact that capital accumulation and labour have had upon each other.

In the U.K., the Census of Production permits some measurement of both the expansion in manufacturing output and the changes in the use of labour which have accompanied it. During the last half-century a doubling of net output has been achieved with a fifty per cent. increase in the total labour force, equipped with nearly four times as much horsepower, and containing more than three times as many "administrative, technical, and clerical employees." The first three discoveries should be neither surprising nor very exciting, except possibly that they can be measured at all over so long a period, but the trend revealed in the occupational structure of the manufacturing labour force has several interesting implications and, in fact, has been the source of some speculation about the nature of productivity achievement.

It has been argued that the association between output expansion and the growth of administrative overheads has some significance in that the latter is quite independent of labour productivity and therefore limits the effects of rising

labour productivity on the output of goods per person. In his examinations of the factors which have determined the major differences of labour productivity between industries and countries, Seymour Melman [1] has attempted to establish two hypotheses, among others, about the association between labour and output expansion: that labour productivity is governed by the degree of mechanisation, and that the growth of "administrative overheads" has limited the effect of rising labour productivity on the output of goods per person.

Statements or arguments of this kind mean either very little or a great deal, and thereby involve a serious controversy about the nature of labour productivity—about the distinction between productive and un-productive labour. But even if they are rejected, the phenomenon of rising administrative overheads remains to be explained and its relation to output expansion defined. All the evidence so far collected in both the U.S. and U.K. points to a marked and persistent lengthening in, to borrow a military phrase, the logistical tail of manufacturing industry. In other words, the size of the labour force employed on tasks directly relating to the production and despatch of goods has been a declining proportion of a growing total for over fifty years.

The aims of this paper are three-fold : to measure, in some detail, the growth of administrative overheads in U.K. manufacturing during this century ; to make comparison with the growth of net output and the use of mechanical power ; and to critically examine Seymour Melman's conclusions about their significance. Rather than answer existing questions about administrative overheads, it will tend to ask new ones and only suggest lines along which the answers might be found. In the following section some problems of measurement are described, in the third section the statistical results are analysed, and in the fourth some conclusions are attempted.

Problems of Measurement.

The only official source which provides information about the use of labour and the growth of output on the same basis and over any length of time is the Census of Production. The

[1] *Dynamic Factors in Industrial Productivity*, Blackwell, Oxford, 1956.

Ministry of Labour provides a more frequent industrial analysis of the labour force, but only back as far as the 1920's. The Census of Population provides a much more detailed occupational analysis of the labour force for a very much longer period, but its figures are derived from returns made by households and not from the industrial establishments themselves. In any case they all used differing industrial classifications before 1948. The Census of Production goes back to 1907, is derived from returns made by manufacturing establishments, and provides one simple, fairly consistent, division of total employees into those engaged on tasks directly relating to the production and despatch of goods, and the rest. Between 1907 and 1948 it was taken very infrequently—between 1935 and 1948 not at all—and the extent of the information it collected, not to mention its definition, was continually changing. For the purposes of this paper its scope has been further restricted by the subtraction of Mining and Quarrying, Building and Contracting, and Gas, Electricity, and Water.

In order to use the more accurate and detailed information of the post 1948 Census years, all the variables are measured in terms of the fourteen major divisions of the Standard Industrial Classification (1951). In effect this means that the figures of the pre-1935 Census reports that have been used in this paper—1907 and 1924—had to be re-classified industrially, with varying degree of success. When the absolute values of the variables are measured for 1951, however, it has been possible to expand the number of industrial groups to twenty-one.

The labour force and its sub-division can be measured with a reasonable degree of accuracy in all the census years. Unfortunately this reasonableness is relative, for the occupational analysis only refers to those establishments employing more than ten persons. The figures of net output, or net value added, cover all establishments but do not make any allowance for the depreciation of capital equipment, whilst data on horse-power capacity was not collected in 1935 and its definition changed between 1930 and 1948. In order to arrive at some measure of productivity for purposes of comparison use has been made of the index number of net output calculated

by Mr. K. S. Lomax [1] from Census of Production material. Its values for the major fourteen industrial groups of the Standard Industrial Classification in 1907, 1924 (=100), 1935 and 1951 have been deflated (or inflated) by the rise (or fall) in total employment, and the process repeated for changes in average weekly hours worked. The latter fluctuated quite markedly over the period—55·8 in 1906, 45·8 in 1924, 47·9 in 1935, and 45·6 in 1951—and cannot be ignored. The resulting index of net output per employee-hour is a crude global measure of labour productivity. In the case of horse-power capacity, the absolute values have been calculated directly from the Census for each industrial group and divided by the number of "productive operatives." This gives an even cruder measure of mechanisation but has to be accepted for the lack of anything else.

Labour productivity is defined here as output per employee not output per "productive operative" as it is by Melman. It will be argued that to divide net output by any part of the labour force rather than by the total is not very meaningful and prejudges the argument about productive and un-productive labour which it is part of the purpose of this paper to make explicit. In fact figures of net output and total employment are more comprehensive than those for "productive operatives" so there is an additional statistical objection. There would appear to be no such conceptual and statistical objection to dividing horse-power capacity by "productive operatives" only. Mechanisation of "administrative, technical, and clerical" operations, when it takes place, would be better measured by electricity used, and the figures for h.p. capacity are no more comprehensive than those for "productive operatives."

If all the variables are converted into an index number (1924 = 100) it becomes possible to compare their rates of change between 1907 and 1951. Because a certain amount of juggling has been necessary to arrive at this stage, the results must be treated with caution. For example, the figures for

[1] "Production and Productivity Movements in the U.K. since 1900," London and Cambridge Economic Service, *Times Review of Industry*, June, 1959.

average weekly working hours relate to 1906 not 1907, and those for h.p. capacity to 1930 not to 1935. It has to be assumed that the years 1907, 1924, 1935, and 1951 are typical and good indicators of the general trend of the first half of this century, and the choice of 1924 as the base year is justified by the particular inadequacies of the 1907 Census.

For the measurement of the absolute values of the variables in 1951 a number of alterations are made. Some of the large industrial groups have been further divided to give a total of twenty-one instead of fourteen. The measure of mechanisation is horse-power in use per "productive operative," not horse-power capacity installed, and net output per employee becomes £ net value added per employee (which may not be itself an improvement).

The Transformation in U.K. Manufacturing.

It is essential to examine in some detail the values of administrative overheads, productivity, and mechanisation, because the overall or aggregate figures can give a misleading picture of what has happened in many of the industrial groups. The figures, and any necessary explanation, have been set out at the end of the paper. What follows is a general description and discussion of their evidence.

Total employment in U.K. manufacturing increased by nearly three millions from 1907 to 1951. Most of the expansion took place between the Census years of 1935 and 1948, and predominantly in "new" industries like Chemicals, Motor Vehicles, and Electrical Engineering. Table IA reveals a marked change in the industrial structure of the labour force. Only Textiles and Clothing show any absolute decline. The metal-using industries (Engineering and Shipbuilding, Vehicles, and all Metal Manufactures) rose from 19·6% of total employment in 1907 to 45·3% in 1951. Similar trends have been observed in other advanced industrial countries, and must be considered in the light of long-term employment changes taking place elsewhere in the economy—a decline in Domestic Service, Mining and Quarrying, and Agriculture, and a rise

in the Service Industries, and Government Employment in particular.[1]

In the Final Report of the 1907 Census employment was divided into "wage earners" and "salaried persons." "Wage earners" were defined as those engaged in the manufacturing process, in cleaning, repairing, or maintenance work, in despatch, or in the immediate supervision of operative staff, whether paid wages or salaries. "Salaried persons" were defined as clerks, typists, salesmen, travellers, buyers, and other management staff. By 1951, these categories had come to be called "productive operatives" and "administrative, technical, and clerical employees." "Productive operatives" now include inspectors, and viewers, whilst "administrative, technical, and clerical employees" include works foremen, and those engaged in experimental, development, and technical work. Although the terms used to describe these categories have a strong emotive content, and suggest social status rather than economic distinctions, they are designed to measure, as far as the Census is concerned, two different uses of labour in the manufacturing establishment. On the whole the "white collared" worker does do a different job from the "manual" worker, and the fact that he received a "salary" rather than a "wage," and is called an "employee" rather than an "operative" is the outward and visible sign. It appears that changes in the classification have been designed to take into account additions to the normal staff of a manufacturing establishment, and that an attempt has been made to retain some consistency in the figures.

The general trend, as shown in Table II, is for "administrative, technical, and clerical employees" to increase even where "productive operatives" decline, as in Textiles. There are some exceptions to this rule particularly between 1907 and 1924, but these might well be due to the importance of small firms in the exceptions and the discrepancy in this respect

[1] *International Labour Review*: "World's Working Population: Its Industrial Distribution," May, 1956; "Its Distribution by Status and Occupation," Aug., 1956.
Also the National Bureau of Economic Research Studies: one by Stigler: *Trends in Employment in the Service Industries*, Princeton University Press, Princeton, 1956, and the other by Abramovitz and Eliasberg: *The Growth of Public Employment in Great Britain*, Princeton University Press, Princeton, 1957.

between the reports for those years which is explained in the notes to Table II. If the number of "administrative, technical, and clerical employees" per hundred "productive operatives" is calculated—Melman's A/P ratio—another useful index of administrative overhead is obtained. The average for total manufacturing rose from 8·7 in 1907, by 11·6 in 1924 and 15·2 in 1935, to 21·0 in 1951. There are only three cases of the A/P ratio actually declining (Leather from 13·4 in 1924 to 13·0 in 1935, Clothing from 11·5 in 1907 to 10·7 in 1935, and Food, Drink, and Tobacco from 26·1 in 1935 to 24·1 in 1951) but its rate of growth varies considerably between the industrial groups. Furthermore there is a remarkable spread in its absolute values, with Textiles and Chemicals at the two extremes throughout the period (in 1907 Textiles was 3·5 and Chemicals 16·2 and by 1951 their respective A/P ratios had become 10·6 and 41·0). There is no evidence that distribution of the benefits, or burdens of administrative overheads has become more equitably distributed as a result of their overall growth. On the other hand there is some evidence that the sufferers or beneficiaries have changed in industrial content. Both in 1907 and 1951 roughly one-third of total "productive-operatives," those in the industries with the highest A/P ratio, worked with roughly half of the "administrative, technical, and clerical employees," and roughly another third, those in the industries with the lowest A/P ratio, worked with roughly another fifth of the "administrative technical, and clerical employees." Since 1907, Food, Leather, and Clothing have dropped out of the "top" group, and Engineering and Ship-building has been promoted.

Similarly, installed h.p. capacity per "productive operative" has shown a steady and persistent rise throughout the period and has been as unequally distributed. The average for total manufacturing rose from 1·3 in 1907, by 2·4 in 1924 and 2·8 in 1930, to 3·4 in 1951. The least mechanised industry in this sense has always been Clothing (0·1 in 1907 and 0·4 in 1951) and the most mechanised industry in this sense has always been Metal Manufactures (4·9 in 1907 and 10·4 in 1951). If anything the spread among the industrial groups is wider than that of the A/P ratio. Only four industries were above

the average in 1907, only three in 1924, with five in 1930 and 1951. The Treatment of Non-metaliferous Mining Products and Chemicals, in addition to Metal Manufactures, were always among them.

Although the general trend in all these variables, and in that of productivity has been upward, it is not possible to begin to speculate about cause and effect until correlations have been calculated on the values for all the industrial groups. To show that productivity advance is generally accompanied by increased mechanisation is not to prove that the latter causes the former, but to show that those industries which have made the best progress in productivity have also had the fastest increases in mechanisation is to go some way to doing so. If a positive correlation between productivity and mechanisation advances was found together with a negative correlation between productivity and administrative overhead advances, there would be some evidence to support Melman's original hypothesis. In other words it would suggest that mechanisation had increased labour-productivity (as output per productive operative in his definition), but the growth of administrative overheads had reduced the benefits to be derived from this process as revealed in a smaller growth of output per employee (his gain per person at work). The "productive operatives" were in fact *the* productive sector of the labour force and in those industries with the fastest growth of "administrative, technical, and clerical employees," they had to support a growing burden of unproductive labour.

This analysis would be quite consistent with the view that the rise in administrative overheads was the result of factors quite independent of those determining output expansion. The only alternative explanation of these correlations would be to discover some additional factor which could explain both the rise in administrative overheads and the associated slowing down of productivity advances. But if the assumption is dropped that only "productive operatives" are responsible for the output gain, there is no reason to expect a negative correlation between administrative overheads and production, and similarly no reason always to expect a positive correlation between mechanisation (horse-power) and productivity. Indeed

in some industries new and improved methods of production (increased output per head) might be expected to lead to a reduction in the use of horse-power, and be the result of, or necessarily imply, a rise in the number of "administrative, technical, and clerical employees."

In Table III the transformation of British Manufacturing in respect of the rates of change of these three variables can be examined from 1907 to 1951. In eight of the fourteen industries output per employee-hour increased faster than administrative overheads, and in only two of the fourteen did output per employee-hour rise faster than installed h.p. capacity per "productive operative." But there is no discernable pattern in the trends, and calculation of correlation co-efficients is not very rewarding. Taking the percentage changes from 1924 to 1951 as in Table III, the coefficient between output per employee-hour and administrative employee per hundred productive operative is only + 0·465 and that between output per employee-hour and installed horse-power per productive operative is only + 0·182. To compare the rate of change over the whole period the difference between the values of the index numbers in Table III for 1907 and 1951 was divided by their average and the correlations calculated on these readings. The results were almost the same : the correlation between output per employee-hour and administrative employees per hundred productive operatives + 0·451 and that between output per employee-hour and installed horse-power installed per productive operative + 0·163.

When the absolute values of the three values are compared for 1951 (Table IV) the resulting correlation co-efficients are not much stronger : that between net output per employee and administrative overheads being + 0·471 and that between net output per employee and horse-power in use per productive operative + 0·266. In 1951 it is also possible to measure, again very crudely, the size of the firm. This has been done in Table IB by calculating the percentage of the total labour force of each industry in establishments employing more than 500 persons. Most establishments are to be found in the range of less than 100 employees, but here the information given by the Census is least detailed and most variable from one industry

to another. The correlation co-efficient between this percentage and the absolute value of the A/P ratio is only + 0·444.

It might well be argued that so small a number of readings should not be expected to reveal any significant correlations. An increase in the number of industrial groups would either make these positive correlations much stronger, or remove them altogether. The evidence from the absolute values of 1951 certainly does not suggest that they would be made much stronger. But Melman in his comprehensive and very detailed study of administrative overheads in U.S. manufacturing from 1899 to 1947, analysing evidence from over 60 industries, reached the same sort of inconclusive results.[1] He found that administrative overheads appeared to be independent of many factors that otherwise differentiate industries, with the possible exception of the absolute values of administrative overheads and the size of the establishment when there appeared to be a negative correlation.

The Significance of Administrative Overheads.

The evidence of the preceding section does not support the hypothesis that administrative overheads are a burden on productivity advances—productivity meaning output per employee. There has been a marked and persistent, if unequally distributed, rise in the proportion of the labour force not employed on tasks directly relating to the production and despatch of goods which is weakly correlated with increases in output per employee-hour. There is some support for the view that this trend is independent of other factors determining output expansion, because it is as marked in declining as it is in expanding industries. These results, disappointingly inconclusive as they are, only raise further questions about the nature and causes of this occupational shift.

The absence of any negative correlation between administrative overheads and output per employees, however, does not counter Melman's argument. This is based, as it was pointed out above in section 2, on a distinction between worker productivity and output gain per person in the labour force.

[1]Seymour Melman : "The Rise of Administrative Overheads in Manufacturing Industries of the United States, 1899-1947," *Oxford Economic Papers* (N.S.), Feb., 1951.

As long as the A/P ratio is increasing, the gap between these measurements will widen. From his examination of the over-all trend in U.K. manufacturing Melman concludes that : "From 1907 to 1948 there has been a persistent shift of man-hours in England's industrial firms towards administrative work. This occurred in manufacturing as a whole and in the motor car industry. This trend would be of interest if only because of its effect on the productivity of the labour force as a whole compared with output per production worker. The growth of A/P takes on special importance, however, in view of its greater rate of increase as compared with worker productivity." [1]

Even where worker or labour productivity rises faster than the A/P ratio, as he argues it did over the same period in U.S., the definitions adopted force the same conclusion : "In sum, we have found that for the labour force as a whole, the shift of man-hours from production to administration has been a brake on the production gain per person in the labour force obtainable from the growth of labour productivity." [2]

Carried to its logical conclusion this means that a point would be reached, if the present trends in the U.K. continue, at which the growth of administrative overheads would be so much faster than that of output per "productive operative" that there would be no gain in output per employee. Or alternatively in a static analysis of two industries with equal net outputs and total labour forces, the industry with the highest A/P ratio must have the highest level of worker productivity or output per "productive operative."

These conclusions only make sense if the implicit assumption upon which they are based is made explicit : that the only meaningful measure of labour productivity is "output per productive operative," and that "productive operatives" or "industrial staff" or "manual labourers" or "wage earners," or whatever else they may be called, *alone* produce the output of manufacturing industry. In the situation of the last fifty years it follows that this section of the labour force has had to carry a growing burden of un-productive administrators. Once this assumption is denied the argument about the widening

[1]Seymour Melman : *Dynamic Factors in Industrial Productivity*, p. 77.
[2]Seymour Melman : *op. cit.*, p. 144.

gap between worker productivity and output gain per person in the labour force becomes a platitude about insignificant definitions, and the important comparison remains that between administrative overheads and output per employee.

In general there is no good case for making a distinction between productive and un-productive labour. It may be possible to distinguish labour which creates income or profits—labour that helps to create a good or service which can be sold —as productive. Such a definition would exclude domestic servants and include the employees of a laundry or a restaurant, exclude the members of a firm's market research department, but include the employees of an independent market research company. In a more popular sense a distinction is often attempted between necessary and unnecessary labour to mean productive and un-productive. If labour can be dismissed without loss to the output of the firm then it can be said to be unnecessary. Unfortunately there are a number of other tests of necessary labour which might be employed, such as necessary for the efficiency or even the welfare of the firm ; and the distinction becomes so difficult to make, there being no reason why "productive operatives" should be excluded from rigorous examination, that in practice its results prove far more generous to "administrators" than intended.[1] On a close analysis of "productive operatives" it is usually discovered that only part of them work directly "at the bench" producing the output of the firm.

In the particular case of manufacturing and the division of its labour force made in the Census of Production, the use of the descriptions productive and unproductive is even less admissable. "Administrative, technical, and clerical employees" or "salaried persons" seem to cover at least four distinct functions in the modern firm : over-all management or decision-making and its supervision, which involves the collection and analysis of data on all the activities of the firm, financial accounting, and communication ; advertising and sales promotion for the output of the firm ; research and technical development in the firm ; and, of special importance to the large firm, the recruitment, payment, welfare, and industrial relations of the labour

[1] *cf.* Marshall : *Principles of Economics*, 8th Edition, Macmillan & Co., London, 1949, Book II, Chapter 3.

force itself. Only the first and last of these could be termed "administrative"; only the first and third could be clearly distinguished as "overheads"; and all of them could be described, in some sense, as "productive." Whether carried on internally, or purchased as a service from an agency, some kind of market study is essential to most manufacturing firms; without technical research and knowledge from some source the firm will never make innovations or keep up with its competitors; and in the absence of a large number of routine clerical operations all the activities of the firm would soon grind to a halt. Employees involved in all four functions can be said, therefore, either to be providing a necessary service to the firm, or directly contributing to the improvement of its output; and in sales promotion or even some aspects of administering the labour force they no more represent "overhead" costs than do many "productive operatives." There is in fact no justification in principle for reserving to the latter the title of "productive." It is true that many of these functions are duplicated within each manufacturing industry, but it does not appear to be one of Melman's objects to attack the "wastes" of private enterprise and competition.

Even with the argument about productive and unproductive labour out of the way, very little can be said about the nature and significance of this change in the use of labour without the evidence of case studies over time in a number of manufacturing firms. On *a priori* grounds clerks and typists would be expected to apply most of the numerical increase in "administrative, technical, and clerical employees," and the first function, that of over-all management and supervision, always to be the most important. The growth of sales or research departments must involve additions to the problems of general organisation, or else their services are purchased from outside; and all departments will have their own supporting staffs, for clerical operatives are common to all aspects of modern management. There is some evidence from other sources to support these generalisations. It has been estimated by cross-reference with the Census of Production, that in 1931 half the "salaried" group would have consisted of clerks and typists.[1] If anything

[1]David Lockwood: *The Blackcoated Worker*, George Allen & Unwin, London, 1958, p. 73.

this ratio should have expanded since then, and according to the Census of Population clerks of all kinds in England and Wales rose from 769,245 in 1911 to 2,132,135 in 1951. If the salary/wage earner classification given in that Census for 1951 is analysed, clerical workers constituted 53·2% of all salary earners. Because this classification includes the large professional group (accountants, lawyers, etc.) there is some justification for believing that the ratio would be higher in Manufacturing alone. The only function for which some separate estimate can be made is technical employees. The Department of Scientific and Industrial Research has made a survey of research and development expenditures and staff in British Manufacturing in 1955.[1] For the fourteen industrial divisions of the Standard Industrial Classification, the total research and development full-time staff was 109,493. This varied from 4·3% of total employment in Chemicals, to 0·10% in Clothing. If the same proportions are assumed for 1951, they only account for a very small number of "administrative, technical, and clerical employees" : at the most just under one-quarter for Vehicles and about one-seventh for Chemicals. Even these are not entirely technical, for the figures include supporting staffs.

The dominance of office-work in all its forms, in terms of numbers at least, does suggest that there is some justification for describing the process as a switch of man-hours to administration—although not necessarily from "production." More clerks and typists could indicate both extensive or intensive enlargements of the functions of management, or merely the blossoming of luxuries which had previously been denied by lack of funds. The addition of new departments such as sales promotion or technical research involves further responsibilities for over-all management, and the changing economic and political environment in which the firm operatives can impose new tasks for management (e.g. P.A.Y.E.). Alternatively increases in the average size of the firm and in the complexities and efficiency of production call for the advantages of speed,

[1]D.S.I.R. : *Estimates of Resources Devoted to Scientific and Engineering Research and Development in British Manufacturing Industry*, 1955, H.M.S.O., London, 1958.

continuity, precision, and specialisation in existing administrative tasks. In practice these factors may be inseparable and reactive,[1] and certainly not measurable by general analysis. As the process seems strong enough to disregard relative productivity levels, it implies that there were some managerial needs which even the most backward sectors of manufacturing felt they could afford. These needs may be connected with the technical advance and output expansion of the period and it might be argued that the wide divergencies in the absolute levels of the A/P ratio were explained by differences in the rate of technical advance among industries; but the rate of technical advance is even more difficult to measure than productivity. Evidence has been collected from a study of twelve Swedish manufacturing firms over the period 1935–50 which is said to show that between two-thirds and four-fifths of the increase in "salaried employees" could be explained by an extensive re-organisation of methods of production, the introduction of new products, and the development of new sales techniques.[2] A similar, if more restricted, illustration can be found in the recent study of the growth of a large steel works in this country.[3]

Although the pervasive increase in the numbers of "administrative, technical, and clerical employees" has been continuous since 1907, there is no reason why it should continue at the same rate and in the same direction. To argue that the rise of "administrative overheads" should be neither surprising nor disturbing, is not to deny that manufacturing firms will find their cost increasingly burdensome. They also try to avoid the cost of "productive operatives," and there need be no other justification for the mechanisation of office-work just as it has been applied on the factory floor. Yet because there are many signs that this process has already begun, it does not follow that the administrative need for office work has declined.

[1] *Cf.* Max Weber: *The Theory of Social and Economic Organisation* (Part I of Wirtschaft und Gesellschaft, ed. T. Parsons), W. Hodge, Edinburgh, 1947, Ch. III and *Max Weber, Essays in Sociology*, eds., Gerth and Mills, Ch. VIII.

[2] Eric Hook: *Salaried Employees and Industrial Transformation*, Industriens Utredningstitut, Stockholm, 1955.

[3] W. H. Scott, A. H. Halsey, J. A. Banks, T. Lupton: *Technical Change and Industrial Relations*, University of Liverpool, Department of Social Science, Liverpool University Press, 1956, Ch. 2.

On the contrary the sharpening of the administrative need may have provoked the mechanisation, and the saving of clerical labour prove to be a welcome, but subsidiary, benefit of the search for more comprehensive and accurate administrative control.[1]

The number of "administrative, technical, and clerical employees" in the labour force in British Manufacturing is remarkable for their uneven distribution between industries, and for how little is known about their nature and significance. It is unfortunate that the emotive term "un-productive" is often used to describe them, and that this distinction is given some support by the classification of the Census of Production. There is obviously a need for more case studies and possibly a change not only in the terms used by the Census of Production, but also in the analysis of the uses of labour which it collects under these headings.

J. Bonner

University College of North Staffordshire.

[1]D.S.I.R. : *Automation*, H.M.S.O., London, 1956, Ch. IIC. D. A. Bell and Pamela Haddy : "Automation for the Accounts Office of a Retail Organisation—A Case Study," *Accounting* Research, Jan, 1958. Pamela Haddy : "Automation in a British Office," *Journal of Industrial Economics*, Feb., 1958.

NOTES TO TABLES.

TABLES IA and II

i. These tables are derived from the Final Reports of the Census of Production for 1907, 1924, 1935 and 1951. They include corrections made in the figures for the earlier years in subsequent Census reports.

ii. All the figures for 1907 include Southern Ireland, for which no correction has been made.

iii. It has been argued that the Census reports for 1907 and 1924 underestimate total employment by as much as 8% and 3% respectively (see A. Maddison : "Output, Employment, and Productivity in British Manufacturing in the last half-century," *Bulletin of the Oxford University Institute of Statistics*, Vol. 17, 1955, p. 384). This is the problem of small and/or excluded firms. The Census reports for 1935 and 1951 make separate, and much more accurate, estimates of employment in establishments of less than 10 persons, and for errors and omissions. Therefore changes in total employment between 1907 and 1924 are not always very accurate in these tables.

iv. The Census reports for 1935 and 1951 only analyse employment by the two main divisions of "productive operatives" and "administrative, technical, and clerical employees" for the large establishments (employing more than 10 persons). In 1907 and 1924 this analysis was supposed to cover all establishments, but figures for the large establishments alone can be obtained for 1924 in a later Census report. Therefore in order to give as long a period as possible on a common basis, the figures in Table II relate to large establishments in all the years except 1907—when there was an underestimate in any case. It follows that total employment in Table II differs from that in Table I, except for 1907.

v. Working proprietors were included in "salaried employees" or "administrative, technical, and clerical employees" in every census year but 1951. Whenever possible the figures for 1951 have been adjusted to conform with this earlier practice.

TABLE I

	A. Total Employment in British Manufacturing 1907–1951 ('000)				B. Percentage of Labour Force in Large Establishments
	1907	1924	1935	1951	1951
1. Treatment of non-metaliferous mining products	218·1	242·1	267·0	308·1	29·6
2. Chemicals and allied trades	141·1	188·3	219·7	387·2	51·3
3. Metal manufactures ...	332·2	352·6	371·7	542·3	65·1
4. Engineering and Ship-building	710·4	832·5	889·8	1767·9	63·4
5. Vehicles	239·2	361·4	513·9	780·8	(65·4)
6. Metal goods n.e.s. ...	240·7	270·4	291·6	453·9	(27·4)
7. Precision instruments, Jewellery, etc.	70·2	90·2	87·1	130·0	34·3
8. Textiles	1260·5	1278·0	1069·4	1009·9	24·9
9. Leather	60·5	65·8	68·5	68·2	4·5
10. Clothing	617·6	612·5	633·1	604·3	17·8
11. Food, Drink and Tobacco	452·1	508·7	633·3	772·8	32·5
12. Wood and Cork ...	191·7	189·8	263·6	284·9	(12·1)
13. Paper and Printing ...	317·4	372·8	434·4	491·0	35·2
14. Other manufactures ...	64·1	108·7	131·5	225·1	47·3
Total	4915·8	5473·8	5874·4	7826·7	36·5

TABLE IB

i. The figures have been calculated from the Final Report of the 1951 Census.

ii. There is some difficulty in making even such a simple analysis as that of establishments employing more than 500 persons. Vehicles and Metal Goods n.e.s. are only detailed for establishments with more than 750 employees. If the figures for 1948 are taken, when the lower limit is taken down to 500, it makes very little difference for Vehicles (65·7) but a considerable difference for Metal Goods n.e.s. Therefore the 65·4 for Vehicles in the Table represents the proportion of the labour force employed in establishments of more than 750 persons in 1951, but the 27·4 for Metal Goods n.e.s. represents the proportion of the labour force employed in establishments of more than 500 persons in 1948. In the case of Wood and Cork information is only given for establishments employing more than 400 persons in both 1948 and 1951, so there is no choice.

iii. "A large establishment" employs more than 500.

TABLE II.—"PRODUCTIVE OPERATIVES" AND "ADMINISTRATIVE, TECHNICAL, AND CLERICAL EMPLOYEES" IN BRITISH MANUFACTURING 1907–1951

	1907			1924			1935			1951		
	P	A	A/P	P	A	A/P	P	A	A/P	P	A	A/P
1. Treatment of non-metaliferous, etc....	205·0	13·1	6·4	213·4	17·5	8·2	235·2	23·3	9·9	257·8	37·8	14·7
2. Chemicals and Allied Trades ...	121·4	19·7	16·2	142·9	35·2	24·6	156·0	50·6	32·4	269·5	110·5	41·0
3. Metal Manufactures	313·7	18·5	5·9	318·3	27·5	8·6	326·5	35·3	10·8	452·9	85·9	19·0
4. Engineering and Shipbuilding	656·9	53·5	8·1	684·8	106·2	15·5	713·6	143·2	20·1	1364·9	372·4	27·3
5. Vehicles	222·4	16·8	7·6	201·5	36·9	11·9	386·4	62·9	13·8	632·8	139·9	22·1
6. Metal goods, n.e.s.	216·3	24·4	11·3	208·0	24·1	11·6	244·3	29·2	12·0	353·7	66·8	18·9
7. Precision Instruments, Jewellery, etc.	61·5	8·7	14·1	65·7	10·5	16·0	58·4	11·7	20·0	94·5	25·0	26·4
8. Textiles	1217·6	42·9	3·5	1198·3	63·8	5·3	978·1	65·4	6·7	904·3	96·0	10·6
9. Leather	53·7	6·8	12·7	48·6	6·5	13·4	52·4	6·8	13·0	51·1	8·7	17·0
10. Clothing	554·1	63·5	11·5	424·7	42·8	10·1	477·0	51·2	10·7	503·4	56·3	11·2
11. Food, Drink and Tobacco	390·6	61·5	15·8	369·6	70·2	19·0	415·6	108·3	26·1	544·6	131·1	24·1
12. Wood and Cork	173·0	18·7	10·8	122·9	14·7	12·0	182·9	23·3	12·7	213·5	33·3	15·6
13. Paper and Printing	280·0	37·4	13·4	290·3	52·4	18·0	331·0	72·0	21·7	367·3	102·1	27·8
14. Other manufactures	57·5	6·6	11·5	88·0	12·9	14·6	105·6	19·7	18·6	177·7	40·8	23·0
TOTAL	4523·7	392·1	8·7	4477·0	521·2	11·6	4633·0	702·9	15·1	6188·0	1306·7	21·0

A = Administrative, technical, and clerical employees (000's).
P = Productive Operatives (000's).
A/P = Administrative, technical, and clerical employees per hundred productive operatives.

TABLE III

i. The index for *a*. has been calculated directly from the A/P values given for all the years in Table II.

ii. The index for *b*. has been calculated from Mr Lomax's index of net output as described in Section 2 of the paper. The figures for total employment in Table I were used to deflate (inflate) his index in order to arrive at an index of net output per employee. The figures for average weekly hours worked were taken from the Ministry of Labour Hours and Earnings Inquiry in October for 1951, 1935 and 1924, and those for 1906 from the Board of Trade Hours and Earnings Inquiry—there being no such information for 1907. Figures of average weekly hours worked only apply to "productive operatives," and it has to be assumed that those for the rest of the labour force move in the same direction and order. In 1906, 1924, and 1935 it has been necessary to re-calculate the industrial analysis of average weekly hours worked in order to fit the S.I.C.

iii. The index of *c*. has been calculated from the values of net horse-power installed per productive operative in all the years taken. Horse-power capacity was measured in 1907, 1924, 1930 and 1951. Once again in 1907 the figures were supposed to relate to all establishments, and in the other years only to the large establishments. In 1951 h.p. capacity was defined as the h.p. of prime movers not coupled to generators plus the h.p. of all electric motors, and 1907 appears to give a similar measure. Unfortunately in 1924 and 1935 h.p. capacity was defined as h.p. of all prime movers plus the h.p. of motors driven by purchased electricity. There is no way of removing this discrepancy—in fact it may not be very important. The figures of "productive operatives" were taken from Table II.

TABLE III

TRANSFORMATION : U.K. MANUFACTURING 1907–1951
(1924 = 100)

a. Administrative, technical, and clerical employees per hundred productive operatives.
b. output per employee hour.
c. installed h.p. capacity per productive operative.

		1907	1924	1935	1951
1. Treatment of non-metaliferous mining products	*a*	78·0	100·0	120·7	179·3
	b	76·5	100·0	137·1	190·0
	c	56·0	100·0	120·0*	164·0
2. Chemicals, etc	*a*	65·8	100·0	131·7	166·7
	b	91·8	100·0	110·3	152·6
	c	54·3	100·0	180·0*	205·7
3. Metal Manufactures	*a*	68·6	100·0	125·6	220·9
	b	81·0	100·0	111·3	132·8
	c	57·0	100·0	115·1*	120·9
4. Engineering and Shipbuilding	*a*	52·3	100·0	129·7	176·1
	b	85·8	100·0	104·8	129·4
	c	38·1	100·0	114·3*	133·3
5. Vehicles	*a*	63·9	100·0	116·0	190·8
	b	46·4	100·0	113·1	162·6
	c	46·7	100·0	100·0*	213·3
6. Metal Goods n.e.s.	*a*	97·4	100·0	103·4	162·9
	b	81·0	100·0	128·4	157·1
	c	31·6	100·0	115·8*	157·9
7. Precision Instruments	*a*	88·1	100·0	125·0	165·0
	b	96·9	100·0	138·9	175·4
	c	25·0	100·0	150·0*	162·5
8. Textile	*a*	66·0	100·0	126·4	200·0
	b	90·1	100·0	126·8	164·3
	c	69·6	100·0	121·7*	121·7
9. Leather	*a*	94·8	100·0	97·0	126·9
	b	74·3	100·0	107·3	114·9
	c	33·3	100·0	113·3*	186·7
10. Clothing	*a*	113·9	100·0	105·9	110·9
	b	89·2	100·0	115·0	115·4
	c	33·3	100·0	100·0*	133·3
11. Food, Drink, and Tobacco	*a*	83·1	100·0	137·4	126·8
	b	71·1	100·0	99·4	116·8
	c	56·2	100·0	118·7*	162·5
12. Wood and Cork	*a*	90·0	100·0	105·8	130·0
	b	71·3	100·0	111·7	121·5
	c	52·9	100·0	111·8*	188·2
13. Paper and Printing	*a*	74·4	100·0	120·5	154·4
	b	64·0	100·0	110·2	163·9
	c	44·4	100·0	161·1*	211·1
14. Other	*a*	78·8	100·0	127·4	152·5
	b	101·3	100·0	141·3	183·7
	c	30·0	100·0	145·0*	180·0
Average for Total Manufactures...	*a*	75·0	100·0	130·2	181·9
	b	75·8	100·0	116·9	151·2
	c	54·2	100·0	116·7*	141·7

* refers to 1930 not 1935.

TABLE IV

i. These figures have been taken directly from the Final Census report for 1951, with h.p. in use rather than h.p. capacity.

ii. The number of groups has been expanded to 21, by dividing up Treatment of Non-Metaliferous Mining Products, Metal Manufactures, Engineering and Shipbuilding, Textiles, and Food, Drink and Tobacco into fairly obvious sub-divisions of the S.I.C.

TABLE IV

THE USE OF LABOUR IN BRITISH MANUFACTURING : 1951

		Total Employment	A/P ratio	Net output per employee £	H.p. in use per productive operative at work
1.	Pottery and Glass	142,729	13·2	555	1·7
2.	Manufactured Fuel and Building Materials	165,416	16·1	738	7·6
3.	Chemicals and Allied Trades ...	387,240	42·5	960	6·7
4.	Ferous Metal Manufactures ...	432,930	17·8	789	8·7
5.	Non-ferous Metal Manufacture	109,344	23·9	923	7·9
6.	Mechanical Engineering ...	936,147	27·5	651	2·9
7.	Electrical Engineering ...	621,343	32·8	593	1·8
8.	Shipbuilding	210,480	12·4	530	2·1
9.	Vehicles	780,799	22·1	634	3·2
10.	Metal goods, n.e.s.	453,916	18·7	657	2·7
11.	Precision Instruments and Jewellery	130,014	26·5	638	1·2
12.	Cotton and Wool	487,450	6·1	638	3·1
13.	Other Textiles	522,498	15·2	626	2·0
14.	Leather and Fur	68,213	16·1	629	2·5
15.	Clothing	604,317	11·1	414	0·3
16.	Food	588,338	23·4	557	2·0
17.	Drink	138,340	26·4	1,094	2·8
18.	Tobacco	46,165	24·9	1,375	0·9
19.	Wood and Cork	284,929	15·7	565	2·9
20.	Paper and Printing	491,011	27·8	852	3·2
21.	Other Manufactures	225,112	22·8	673	3·5
	Total Manufacturing ...	7,826,731	21·1	677	3·3

Stability in Oligopoly

A definitive and comprehensive theory of oligopoly continues to be illusive, but valid and useful conclusions can be reached about phases of the problem. Such theorizing may lack the formal elegance of competitive price theory but the corresponding aspects of policy will be illuminated. Furthermore, each usable interpretation of a part of the whole problem will support another and progress will be made toward the completed theoretical structure. Such is true of the segment to be considered here, the issue as to whether oligopolistic markets are inherently unstable in the absence of effective collusion.[1] Once that is settled, progress on other phases of the larger problem will be expedited.

For what is to be done here these other unsettled issues concerning the behaviour of oligopolistic markets are set aside. The question as to whether an oligopoly can be said to have a stable equilibrium is to be explored separately and prior to examining the welfare properties of equilibria if there be such in these markets. Nor is it necessary to consider the reaction of an oligopoly to such exogenous developments as movements of factor prices or of demand for the industry's product. Indeed, in this piece by piece attack on the oligopoly problem, it is contended that, in contrast to the practice in the theory of competition, the equilibrium problem under stable external conditions should not only be analyzed first and separately, but also as a problem different in kind. The concern here is with inter-firm rivalry (within an " entry limit ") under given factor-price and demand conditions and of that problem only the stability tendency, and not the cost-price output relationships in a condition of stability, is to be explored.

[1]One need not belabor the place the instability view of oligopoly, or the view that if stability exists it is maintained by collusion—a frequent feature of American anti-trust indictments—has occupied in the literature. But it is significant that in a recent attempt to show that equilibrium is possible even if each firm has different attitude towards uncertainty, Kaysen concluded that what he found is only a " quasi-equilibrium " or a condition of relative stability of " going values ", for he is not certain " that a small random disturbance which pushes prices away from those going values will be followed by a return to them." Carl Kaysen, " Dynamic Aspects of Oligopoly Price Theory ", *American Economic Review Papers and Proceedings*, Vol. 42, May 1952, p. 209.

Even the limited aspect of oligopoly theory to be studied here cannot be solved unless the perennial stumbling block of conjectural interdependence can be removed or a satisfactory route be found around it. The circularity of expectations as to the effect of a competitive move by one firm stems often implicitly from the assumption that the firms are alike in relevant attributes—as if they had been created *de novo*—and had available only one " paramater of action,"[1] a change of price. Then there is no room for independent action without disastrous results on joint profits, and no plausible way whereby under neo-classical type assumptions a definitive equilibrium can be reached without overt collusion. The issue that arises, then, is whether one finds, empirically in markets where sellers are few, the type of structural situation just alluded to and, correspondingly reliance on price competition or on other forms of rivalry that pose equivalent problems of " agreement " via conjectural interdependence. It will not be surprising, therefore, to find that the analysis here will turn in large part on alternative assumptions concerning the relevant structural features of firms and of the market and the forms of rivalry used.

The " Stability Zone " Concept

Partially because of these assumptions and because their significance stems to a considerable degree from historical-institutional considerations, the term equilibrium will be avoided here. In the literature it has acquired connotations foreign to an analysis that does not find the conditions of an equilibrium to be rooted directly in technological and utility functions, as in a competitive equilibrium. And for that reason and also for a reason to be pointed out in a moment, the term "stability zone " will be used here. But as in equilibrium, a stability zone refers to a relationship among sellers' market shares and their respective prices and costs. For these relationships to be in that zone, it is not necessary that all firms have the same costs or prices. Differences must reflect a balance among the respective firms' volumes and buyers' preferences for the various firms' " products ". But neither the market shares, the prices, nor the gross margins over direct costs, of the rival firms need bear a

[1]Tibor Scitovsky's term (*Welfare and Competition*, Irwin, Homewood, Illinois, p.9).

precisely fixed relationship to each other. Such a stability zone can encompass a slow drift, particularly in market shares, and perhaps in the level of prices compared to costs.[1] Movements of that sort are not disruptive because an oligopoly has strong centripetal properties for reasons to be developed in a moment.

If movement were to occur beyond the limits of this zone some firm adversely affected which has the power to do so would decide to act independently in order to defend its longer term competitive position which is the same thing as the future value of the firm. Then cumulative adjustment is likely. But the danger of that, and its costs, to the initiator as well as to rivals is one of the sources of the centripetal tendency referred to above. Only when the future value of the enterprise is expected to be enhanced sufficiently would a positive, not defensive, move be made to disrupt the stability zone.

Stabilizing Effect of Firms' Attributes Acquired over Time

The first evidence of the stability tendency in oligopoly is found in the process by which the market shares and cost-price relationships settle in the stability zone. The method used in competitive theory, which is essentially that of analyzing this process as if all firms were extablished *de novo*. seems inappropriate here. In part the reason lies in a major characteristic of the industries in which oligopoly is important ; typically they have a high fixed capital-output ratio.[2] For this reason the planning horizon of the firm is long and, by the same token, what the firm is or plans to become cannot be cut loose from its history. In addition, lacking a theory for a definitive appraisal of what forces are relevant to, and their influence on, the equilibrium aspects of oligopoly, pragmatism forces emphasis on the historical evolution of an oligopoly's structure and internal relationships. That becomes more persuasive when one notes important characteristics of these markets and of the firms that are in them.

[1]The analogy of this equilibrium, or condition of stability, with J. S. Bain's " persistent conditions ", " that are usually stable and slowly changing through time ", is evident. (*Barriers to New Competition*, University of California Press, Berkeley, 1956, pp.7, 18.)

[2]One would deduce from competitive theory that, where this technological feature is important in a competitively-structured industry, socially optimal performance could be expected. But such results have not been demonstrated, empirically.

Instead of these being a dominating homogeneity of process, of products, of buyers' wants and of sellers' abilities, there is almost without exception a significant, and often substantial degree of heterogeneity. This is as true of industrial products as of consumers goods and applies nearly always to semi-manufactured items. Each firm comes, through its history, to " fit into " the market, and into the roles played therein by rivals. By a combination of design and accident, each firm develops some diversity in attributes which balance a diversity in factor and product markets and also the characteristic that rivals have or acquire simultaneously. Hetergeneous oligopoly is the rule, but diverse in more ways than just product differentiation, pure oligopoly is a curiosity.

Each firm evolves into a " structural position " of value to itself. That position is a composite of its physical assets, its organization, its product lines, its location, and its relations with suppliers and customers. Rarely would it be wise for it to attempt to change those characteristics quickly, even if it could. What it has become, conditions both what it is wise for it to do and what it can do.

Perhaps the following analogy will help. Pure oligopoly, as visualized in theory, involves not merely standardization of the product but also buyers' complete indifference among potential suppliers. It includes one aspect of an " open market " ; that is, that each transaction is completely autonomous from all past and future transactions between a given seller and a given buyer. Such a market is as if two smooth pieces of wood, one representing buyers and the other sellers, were set side by side and a person could move one along the other with the slightest effort. But as an oligopoly evolves, with there being a significant degree of uniqueness of firms and customers, of " products ", of locations, etc., each piece of wood, to pursue the analogy, acquires indentations and excrescences, and in time those on each piece " grow " in ways that fit the other. Then more than a slight effort is required to slide one along the other ; the effort must be enough to break the fit between the various indentations and the excrescences. Structural imperfections, discovered or initiated historically, continue their influence in the years that follow.

Along with properties of firms *vis a vis* characteristics of product demand (and possibly also of factor markets) there evolve practices which have a stabilizing influence. Their general affect is to escape from an " open market " situation entirely to one of enduring supplier-customer relations of value to, and preferred by, both buyers and sellers.[1] In consumer goods industries these consist not merely of developing consumers' preferences for brands[2] but also, and often of more importance, of relations with the distributive trades. In producer goods, including non-agricultural raw materials and partially manufactured goods of all sorts—the area where oligopoly is most significant—what tends to develop is a " best favored nation " arrangement. Customers expect their supply needs to be taken care of so as not to make their own operations inefficient and to be given as low a price as is quoted (at least openly) to anyone else. On the other side, the supplier expects to be given the chance to meet any lower price of a potential alternative source. (In a weak market a buyer with excellent alternatives may be able to obtain a secret concession that is not extended to even " best favored " customers). All such practices have the effect of minimizing the gain of volume that any one firm can expect from initiating price competition.

For this, and a reason yet to be developed, overt price competition is rare or absent and rivalry takes forms that do not shift market shares rapidly. This is true of the method so much emphasized, advertising, but also applies to " better service " to business buyers, small " improvements " in products, development of reciprocal arrangements between firms, and so forth. For each successful move of these sorts by a rival, ordinarily there is a counter move of some potential effectiveness. Ordinarily both the initiator and the reactor act independently,

[1]Hetergeneity is viewed here as more pervasive (including non-consumer goods) and less adapted to static and more to historical and dynamic analysis than one sees in the writings of E. H. Chamberlin, for example.

[2]When one says differentiation is " of value " to the consumers, that does not belie some adverse welfare effects although spelling them out, without relying on some " ideal type " concept of the consumer and a view of consumer behaviour of questionable validity, has not been done very satisfactorily. The point here is that the consumers' conduct indicates a preference to rely on brands and particular sellers rather than to re-appraise alternatives before each purchase.

that is, without being restrained by expectations as to rival's conduct. While imitation usually lags, the give and take among rivals means that the " fit " between them is nearly always in a degree of flux, rarely does any movement of this sort disrupt the stability zone. Nor is that intended (except under two circumstances yet to be outlined) in part because the penalty for doing so is so serious, as will be seen in a moment.

Over the years not only do rival firms evolve so as to fit into each other and into the markets, and develop stabilizing practices, but also they have repeated experience with each other's reactions. In part these are observations of actions and reactions that constitute rational behavior in the structural situations in which the various firms find themselves and the values of relevant variables at particular times. To that extent one can be optimistic about the chances of developing generalisations about behavior. Indeed, the author is inclined to the conclusion that the unique person in a responsible decision-making position is relatively rare and that most men who are likely to occupy those posts will make about the same choices in a given structural situation and where faced by a given set of data.[1] That is what was meant earlier when it was said that management of each firm comes to have a quite clear view of it can do or of what it would be wise for it to do, and simultaneously, a similar judgment as to rivals' situations. While much of this, it is contended, is deduced by such officials from the structural situation in which various firms find themselves, both as to their positions in the market and the respective firms' attributes, it also is rooted in experience.

There is a learning, as well as a deductive basis, for anticipating conduct of rivals. Given those expectations and the firm's position in the structure of the market, its choice among the feasible alternatives before it contains a logic of the sort economists have always sought.

A major consequence of such decision making, and one taught by both deduction and experience is the penalty to be expected from taking actions beyond the limits of the stability zone. The exception would be when disturbance is done intentionally and

[1]For an opposite view see Fritz Machlup *Economics of Sellers' Competition* (The John Hopkins University Press, Baltimore, 1950), p. 420.

with foresight as part of a major program which promises to enhance the long run rate of profit of the disturber. Such cases, termed " planned disturbances " here, are set aside for the present but it should be noted that the title given them connotes ample preparation for reactions of rivals and careful weighing of the outcome. In contrast, in the day-to-day, smaller-scale and non-price rivalry, each firm is tempted to push toward the boundary beyond which rivals would, in self defense, react strongly and in a fashion that quite likely would reduce aggregate gross and net profit of the industry. Being unprepared to stop such reactions, nor wishing its consequences, going that far is termed " unplanned ".

Nevertheless there is a tendency to push toward that brink for a variety of reasons. The most pressing reason is to put unutilized capacity to work whether the idle capacity is installed or expected because the feasible increment of expansion is large. In an oligopoly each firm may have a margin of excess capacity most of the time. On the other hand only some firms may have it, or their percentage of idle capacity may be much higher than the average. This could arise because of discontinuities involved in optimal increments of expansion or because of the firm's loss of volume relative to capacity arising from the slow adverse drift of market shares referred to earlier. The industry as a whole may have misjudged the growth of demand, or new plants may have been built because of shifts in the advantage of locations or of new technology, without the old plants having been dismantled. Finally, of course, excess capacity may be the result of a cyclical decline of demand for the industry's product, a situation that enters this analysis only indirectly.

But, as was pointed out above, each firm's efforts to do better that push it toward the brink, tend to be restrained by rivals' counter-actions. That would not be true, and the stability zone be preserved, were the firm to engage in overt, or extended covert, price competition. But that would not be the method chosen for the several reasons given earlier. Rather it is the varied forms of non-price rivalry by which each firm pushes as if to move to the brink but which other sellers, also by their non-price competition, ordinarily check before movement of market share goes beyond the boundary of the stability zone. If

not, the transition to a new stability zone is usually so gradual that it can be made gracefully even though market shares in the end are quite different. If demand for the industry's product is growing, the usual situation, the chance of the stability zone being disrupted is nominal.

Penalty for Disturbance :

Cost and Demand Characteristics of Most Oligopolistic Industries

The severity of the penalty for " unplanned " movement beyond the boundary of the stability zone stems from certain characteristics of most major lines of activity in which oligopolistic structures are found.[1] These are : a quite, or even extremely, inelastic demand in the short run ; a rather wide variability of that demand cyclically or even more unsystematically ; a low ratio of variable costs to price ; and a necessarily long planning horizon for the firm. The latter stems in part from the fixed capital to output ratio but may also reflect other assets, some of them intangible. Not all of these characteristics are necessarily found in a line of activity which is in fact structured oligopolistically, and some industries may have that structure without that fact being reflective of these influences at all. Of the four types of influences listed, number one, taken in connection with number two, and number four require the most explanation.

First, why do most large-volume oligopolistic industries have the first of these characteristics ? Typically, these are the non-agricultural raw, or semi-manufactured goods industries. Those that do make finished consumer goods either make durables or such non-durable goods as gasoline that are used in connection with the durable to provide the service the household consumes. Consequently, the demand for them is either derived from the demand for final goods or is one element of a joint demand. All such characteristics contribute to inelasticity of demand and also, except for the non-durable goods associated in use with a durable, to wide cyclical variability of demand.

[1]It is tempting to pause on the question as to whether these conditions are themselves an explanation of the evolution of an oligopolistic structure. Since they are related to the potentiality of, or bitter experience with, wide price declines of a degree that could cause financial embarrassment to firms unless they are unusually strong, it is possible that by intention, or survival ability, firms have become larger than the *minimum* optimal size.

The price inelasticity can be appraised as follows. Assume the demand for the final goods or service to have unit elasticity. Then were a ten per cent. decline in its price to occur with an accompanying ten per cent. increase of purchases, and were the dollar and cent amount of that decline passed through to the price for a component of the finished goods, a much higher percentage drop in the price of the component would be necessary in order for the industry at that level to be able to sell 110 per cent. of its previous volume. A similar analysis could be made of the elasticity of demand for one of the several goods involved in rendering a service (such as gasoline as part of the joint demand that provides automotive transportation). Obviously, such a dollar-and-cent pass through would occur only if the supply and demand situation for each other component of the final product, or of the service rendered by a durable good, were such that no one of them would have to share the price decline.

Under those conditions the elasticity of demand for the component of a final good, or of one element of a composite, such as a service rendered by several commodities being used together, can be computed easily, *ceteris paribus*. It is

$$\mu_c = \mu_g \frac{V_c}{P_g} \text{ in which}$$

μ_c is elasticity of demand for the components of a final good or of components of a family of goods or of a service which have a joint demand relationship, μ_g is the elasticity of demand for the final good or service. V_c is value of the component in a unit of the final good or service, and P_g is the typical price of the final good or service in the zone under consideration.

If $\mu_g = 1{\cdot}00$, and $\frac{V_c}{P_g} = {\cdot}4$ then $\mu_c = {\cdot}4$

The significance of inelasticity of demand for the components of final goods or services should be examined in connection with the use of price competition as a means of increasing market share. But this will be explored with reference to a situation in which both factor price and the demand for the aggregate output of the industry are assumed to be stable. That isolates the struggle over market shares when the motives are " internal ", that is

each firm is attempting to improve its relative cost, revenue and profit position, as distinct from the adaptation to new external conditions. Leaving out adaptation to movements of demand also minimizes the chance that the price elasticity deduced by the method suggested above will hold. It rests on the assumption that prices and costs of other components of the final good or service will not be changed so that the dollar-and-cents amount of the price change of the final good will be passed back to earlier-stage transactions.

Under the circumstances, the firm that disturbs the equilibrium must expect strong reactions from rivals to the degree that is attracts customers from them for it forces them to have excess capacity. In addition, the loss of customer connections forecasts lower profits in the future for those firms either as they incur costs to win back customers, or suffer fairly long-period excess capacity. The likelihood is, therefore, that they will promptly match the price cut. Indeed, the initiator can be certain of this reaction and also that potential new customers will give their regular suppliers a chance to match the cut. Consequently, the increase of volume that its initiator can expect will be only that warranted by the relatively inelastic industry demand curve for the component. The resulting increase of volume, it was pointed out above, will be much less in per cent than the drop in price (assuming unit or lower elasticity of demand for the final product).

The prospects that the price decline would be at least as severe as that which would occur if the industry were competitively-structured, makes the elasticity of marginal cost curves particularly significant. If they are highly inelastic in the volume range that is relevant, the check on the decline of prices stemming from reduction of output will be slight. But most of the empirical evidence points toward only modestly rising, or even approximately horizontal marginal (money) cost curves in most manufacturing at least in the relevant range.[1] Hence the shape

[1]The types of production in which the sharply rising marginal cost curves (perhaps that start at zero at output rates not far below capacity) are most apt to be found is where processing occurs in giant container-type units used in a continuous process operation. But most plants have several such units, often each below optimal scale, and adaptations can be made with relatively little more disadvantage than in industries where small increments of output can be added or subtracted. Even if that cannot be done, more costs can be made variable by operating at capacity for a while and then closing down until inventories are reduced.

of the money marginal cost curves' shape appear to ease the impact of an unplanned disturbance, but that is not the end.

The reaction of rivals is not governed solely by immediate cost and revenue considerations because of the length of their planning horizon. Where loss of volume affects the future profitable use of "sunk capital", or destroys valuable trade connections, short term operations take on a different hue. Loss of customer connections can mean reduced future earnings even after the adjustment process for the industry as a whole has been completed. Loss of revenue above cash cost reduces the present value of the enterprise and, in addition, the position of the equity owners may become still worse if insolvency is forced. Whether such "costs" are viewed as "subtractions" from money marginal costs, or "additions" to immediate revenue gained by reducing prices, rivals are under great pressure to match price cuts or even more than to match them. Were the contest over shares, via price cuts, to reach the point that prices fell below money marginal costs (or average variable costs in some parts of the volume range) the cut-throat competition stage will have been reached. While such a stage might pay in a "planned disturbance" as the means of expediting the readjustment, one can hardly see how the initiator gains from such a consequence of an unplanned disturbance.

Now introduce the second characteristic of many of these industries, that of wide cyclical movements of demand. Doing so does not mean that the industry's capacity or its cost-price margins are out of line with long-run influences. But the mode of reaction to the short-term exogenous developments, particularly a fall of demand, can disturb the longer-run balance. Were a firm, in its desire to utilize its (now enlarged) excess capacity, to go beyond the stability zone (indeed that zone may shrink !) he would find an aggravated tendency to meet his efforts. In addition to all of the reasons given above, there is added the impact of the enlarged margin of excess capacity of rivals. Consequently, unless this happens to be fortuitous occasion for a planned disturbance—an unlikely coincidence—the penalty for disturbing the stability zone [1] is more certain

[1] The effect of falling volume on unit costs and gross and net margins per unit is not relevant ; that aspect of the stability zone was defined in terms of a longer-run average experience.

and more severe. The logic of the situation, and of repetitive experience, teach this and underscore the fact of limits to the stability zone. Add what merit there is in the kinked demand curve—the greater difficulty of raising than of reducing prices—repetitive cyclical experience underscores the premium on avoiding actions that become an unplanned disturbance of the stability zone. But it is not behaviour during cyclical movements that is of concern here ; it is the re-inforcing effect of their likelihood on the choice of methods of rivalry that is relevant.

If one adds together unelastic demand, cyclically unstable demand, and the strength of the resistance to abrupt loss of market share, and more so of absolute volume, the penalty for disturbing the stability zone appears very large and certain. That does not mean that rivalry is absent. Rather the methods adopted work slowly and are blunted effectively in most cases by rivals as was outlined in detail earlier.

To complete the account, it should be recognized that a firm would knowingly disturb the equilibrium under two circumstances. One is where, a creeping loss of market share at some stage results in a threat to the firm's solvency. Then, as an act of desperation, it might open the price war in the hope of adding enough to its receipts above cash expenses to ease its emergency. Even then it would often find it better to sell its business at a fraction of its book value, because of the small chance of more than a temporary surcease to its problems ; it is by definition not a strong competitor and temporary financial relief will not enable it to become one. The other possible disturbance would be that by a strong firm which expects to augment its return on capital in the long run, including an uncertainty premium, enough to warrant the vigorous and perhaps prolonged battle to substantially change its market share.

Such a " planned disturbance " is a major undertaking and one for which careful preparations are made for " battle ". It is quite unlike the situation in which a firm may find itself when it had been testing the limits of the stability zone and had gone so far that it opened a battle unintentionally. Even in a planned disturbance, an overt price cut is by no means the usual method of achieving the initiator's ends. That method is apt to be so

costly that slower-working means would be preferred. These are sales effort, reciprocal arrangements if feasible, better " service ", improved product, etc. The method chosen depends on the facts in the case. Were a selling price reduction, relative to factor prices, to be the method it quite likely would be less disturbing if the firm were to fail to respond on the occasion of an advance of factor prices or of demand than to reduce prices relative to direct costs. Indeed, the occasion of such an exogenous development may be selected for instituting a planned disturbance.

Some Concluding Comments

Otherwise the centripetal tendency of an oligopoly, with variations in market shares and in cost-price margins tending to remain within a stability zone, leads one to expect that reactions to factor prices and product demand movements will be such as to avoid disturbing the balance materially. That forecasts a tendency to pass through major factor price movements but to react to demand variations primarily in volume rate unless the extent and duration of the demand change and, on the downside, its erosion by indirect price means, forces someone to make a major and overt price reduction. But further pursuit of the topic of price-volume adaptation to developments from outside the oligopoly is beyond this paper. It is a line of investigation that can be analyzed with more confidence and accuracy once the strength of the stability tendency in oligopoly is recognized.

In the author's judgment what has been said about the stability zone differs from the " full-cost " doctrine in all but one regard. It carries none of the business mores overtone that prices ought to bear a certain relationship to costs. Nor does it rest on the view that businessmen have some set notion of what gross or net profit margins are enduring. It does not connote that prices have no relation to demand nor to the appropriate concept of marginal cost, with user cost given due weight. Rather the gross and net margins (as well as market shares, differences in " products " and in channels of sales) are those " hammered out ", and in a sense tested, over a substantial period of time. Both cost and demand considerations have affected this historical process. Then as to the focus of the analysis, instead of emphasizing the individual firm's " setting " prices on the basis

of its own costs, concern here is with the industry and the relationship among the volumes, " products ", prices, and cost-price margins of the firms in the group. Only in a loose meaning, compared to the usage in price theory, could one term as competitive the process where—by this quite stable set of relationships is achieved. It does emerge from rivalry as it has been, but since that, in an oligopoly, need not be competitive in the strict sense, one cannot deduce therefrom information about the consequent relation of prices to costs. That fundamental property of an equilibrium is not dealt with. But what is in common with at least a part of the full-cost literature is the emphasis on the stability of the relation (or perhaps a slow drift as will be pointed out below) between the average level of costs over a period of time and the average level of at least the overt prices. That is certainly true under stable (or steadily growing) demand, and factor prices changes, if substantial, tend to be " passed through ". Adaptation to cyclical demand movements, to complete the account but not of direct concern here, tends not to be in overt price changes.

Two final comments are in order ; the first of which is of a methodological sort. The " structural situation " of the firm, and the influence of that on its logical conduct in a given situation, are in a sense institutional. They do not stem either directly or in a clearly identifiable fashion from current technological and demand functions. But these institutions themselves reflect past attempts to take advantage of cost and demand conditions of earlier years, modified by mistakes and luck. Furthermore,those past decisions whose shadow limit the present feasible alternatives. The choices made also reflect the quality of information available to firms and their aversion to risk. There is not a crippling loss of generality from introducing such historical-institutional consideration for the generalizations reached have to do with the characteristics of firms and markets that emerge from this process and the effect of these structural features on their behavior. It is not possible yet to spell out fully the functional relationships governing that behavior in this milieu. It is clear that because little or no advantage is gained from use of methods of rivalry that would, if not countered, shift market shares rapidly, the competitive forms actually adopted are

those which promise individual advantage. These are of the sort that do not, in part because of counteraction by rivals, shift volume rapidly or disrupt the stability zone. Preference for such courses of action is reinforced by cost and demand characteristics of most oligopolistic industries which turn price competition into a certain penalty for all firms except where one of them is undertaking a " planned disturbance ".

The final comment has to do with collusion. The argument of this paper is to the effect that stability in an oligopoly is possible and probable without overt collusion, as the consequence of rational and largely independent decisions made by each firm. Where, however, the common interests of the sellers are reinforced by the somewhat elastic glue of collusion, forming the agreement and adherence to it are made easier by the centripetal tendency demonstrated here. By the same token, the stronger the centripetal tendency, the less need for a formal collusion from the point of view of the firms. In any circumstances it can reduce remaining doubt. Doing so is less urgent with respect to maintenance of the stability zone under given external conditions, the topic explored here, than with respect to the course of action or inaction in the event of a fall of demand relative to capacity.

Richard B. Heflebower

Northwestern University

Wages, Social Income, and the Family

I.

The concept of the dependent family has featured prominently in the discussion of wage problems. It is not the purpose of this article to discuss whether the wage a man receives ought normally to be sufficient to provide for the needs of a family : its purpose is rather to see how far in fact the present wage system does provide for family needs, and how the standard of living of the dependent family compares with that of other households in the community. In particular, an attempt is made to show the influence of contemporary social policy in raising the living standards of the dependent family.

It is sometimes argued that the dependent family is no longer typical, if indeed it ever were. In fact, in 1951, only one wage earner in six was the sole-earner in a household that included dependent children.[1] This figure, however, is probably nearer one in four if we consider only male wage-earners. Moreover, in the largest group of households, namely those with both dependent children and secondary wage-earners, there will be many cases where the secondary earners are young people who make little or no contribution from their earnings to the family purse.

Thus the dependent family with a single wage-earner, or a close approximation to it, is not so rare as is sometimes supposed, and is sufficiently common to merit serious attention, especially in view of the fact that it represents a stage through which most households must pass, and a very important and formative stage at that.

Under the present wage system, where a man's income is related to the market value of his labour rather than to his needs, there is inevitably a big difference in living standards between the childless household and the dependent family, even if the rise in the level of productivity now makes it possible to pay most men a wage sufficient for the support of a family.

[1]See, T. Brennan, "Household Structure and Family Income," *Scottish Journal of Political Economy*, 1955.

Lady Gertrude Williams has argued that the wage system has already been modified or supplemented to a significant extent, and that "members of the community—individuals or families—no longer have to adjust their standards of living to their earnings" for "society provides a 'social income' that bears no relationship to earning power, and, unlike wages, has no difficulty in adjusting itself to varying needs." She does not claim that "this social income now puts the family man on the same footing as the bachelor or the childless couple" but suggests that it could do so if we wished.[1]

II.

In our attempt to show how far the wage system caters for the needs of the dependent family, we must estimate the income it requires. We can take as our starting point Rowntree's "human needs" standard, and calculate the income required to provide this standard at the present time. Rowntree and Lavers made such a calculation in 1950.[2] This was that a family of five (man, wife and three children) needed an income, after all deductions, of £5 a week plus whatever payments were being made for accommodation. In 1954, the average rent paid by families with an income between £6 and £8 a week was 14*s.* 1½*d.* a week.[3] This can be reduced to the equivalent of 12*s.* 4*d.* at 1950 rents by the use of the appropriate price relative. This gives the total "take-home pay" required by the family of five in 1950 as £5 12*s.* 4*d.* This can be converted to the equivalent of £8 9*s.* at 1959 prices by the use of an appropriate price index.[4] To this must be added 9*s.* 11*d.* to cover National Insurance contributions. A man with three children would receive, however, 18*s.* in Family Allowances, so that he could just manage to support his family at the

[1] G. Williams, "The Myth of Fair Wages," *Economic Journal*, 1956.

[2] B. S. Rowntree and G. R. Lavers, *Poverty and the Welfare State*, London, 1950, p. 24.

[3] H. F. Lydall, *British Incomes and Savings*, Oxford, 1955, p. 190, Table 97.

[4] The price index used for this purpose is one calculated by a colleague, Mr. K. R. Hutchings in an unpublished paper, "Prices, Taxation and the Family." It is calculated from the officially published price relatives but uses the weights appropriate to the particular household size and income range instead of those appropriate to the whole sample, as in the official index of retail prices.

"human needs" standard if his wage, before deductions, was £8 1*s*. a week.

It is not easy to say how far such a wage is being currently paid to adult male wage-earners in Britain, for there is more information available about basic rates than about actual earnings. Moreover, the difference between earnings and wage rates is not simply a matter of overtime. Differences arise from additional payments for shift working, from schemes for payment by results, and because with a shortage of labour employers are forced to offer more than the basic rates agreed with the unions. Nevertheless, there are probably some men in most industries who have no opportunity of earning more than the basic time rates determined by collective bargaining or statutory wage-fixing machinery. Moreover, this situation is likely to be more common in the case of the lowest grade of labour than it is for the higher grades. There is no data published to show how many men are actually earning less than £8 1*s*. a week, or even how many are employed in industries and categories where the minimum rate is below this figure. All that can be said is that of the agreements recorded by the Ministry of Labour in April, 1959, the basic rate for the lowest grade of labour was below this figure in 103 cases and equal to or greater than this figure in 95.[1]

III.

Table I shows the disposable income of representative households after payment of income tax and the receipt of Family Allowances. From this table it is possible to obtain some idea of the relative standards of living of childless couples and dependent families. For this purpose, a simple comparison of income per head would not be appropriate. There may be "economies of scale" in household management, and in some respects the needs of children may be met more cheaply than the corresponding needs of adults. In addition, a family may impose costs which cannot be measured, in the sacrifice of leisure, particularly by the mother. This may mean a lower standard of living for the dependent family, though it may not be capable of demonstration in any statistical data.

[1]Ministry of Labour, *Time Rates of Wages and Hours of Labour*, 1959.

To make a rough comparison of living standards, one would have to make some estimate of the relative cost of maintaining an adult and a child. It is probably not too far wide of the mark to put the cost of maintaining a child at half the cost of maintaining an adult.[1]

TABLE I

DISPOSABLE INCOME OF REPRESENTATIVE HOUSEHOLDS

	Married couple with number of children as below :					
	0	2	4	0	2	4
Wage per annum	£450	£450	£450	£650	£650	£650
Family Allowances * ...	—	£20	£72	—	£20	£72
Income Tax*	£16	—	—	£58	£6	—
Disposable Income	£434 (100)	£470 (108)	£522 (120)	£584 (100)	£664 (114)	£722 (124)

* Rates current during financial year 1959/60.

Family Allowances and income tax reliefs are only two of the benefits conferred upon families by contemporary social policy. In Table II, an attempt is made to take account also of the other principal benefits provided by the state which are divisible and which are received as of right.[2] These are the Welfare Foods, school milk, and the National Health Service. The value of the Welfare Food and school milk subsidies, averaged over the period of fifteen years that a child is normally dependent, is about £5 a year. The allocation of the cost of the National Health Service to particular categories of users

[1]Thus in 1950, Rowntree and Lavers, estimated the income needed by the family of five as being 75% higher than that required by the childless couple. This figure was exclusive of rent, which is perhaps one of the major items that does not increase in proportion to family size. On the other hand, this very fact reflects a lower standard of housing (and therefore of living) for the dependent family.

[2]In this article, housing subsidies are excluded. They are not received as "of right" in the sense that they are available to all in clearly defined circumstances who wish to accept them. The exclusion of the value of educational services provided freely by the State is based on other grounds. The family benefits from having free education available for its children, but this is a special service required by families with children. To include the value of these services in the "social income" received by the family would involve a form of "double counting" if used in a comparison of living standards.

is not practicable. Some age-groups make bigger demands on it than others. Families with young children may make heavy demands on their general practitioners, whilst the older age-groups make the heaviest demands on the expensive specialist and hospital services. In Table II, the cost has been allocated on a simple *per capita* basis at about £12. It seems unlikely that this will overestimate the value of the service to families with children : the opposite is more likely.

TABLE II
TOTAL REAL INCOME OF REPRESENTATIVE HOUSEHOLDS

	Married couple with number of children as below :					
Children	0	2	4	0	2	4
Wage per annum	£450	£450	£450	£650	£650	£650
Disposable Income ... ... (from Table I)	£434 (100)	£470 (108)	£614 (120)	£584 (100)	£664 (114)	£722 (124)
Social Income *	£24	£58	£92	£24	£58	£92
Total Real Income † ...	£458 (100)	£528 (115)	£614 (133)	£608 (100)	£722 (120)	£814 (136)

* Value of welfare food and school milk subsidies plus value of National Health Service benefits.
† Disposable Income plus Social Income.

Considering Tables I and II together, it appears that the results of present social policy include the equalisation of living standards between different sized households to a very limited extent only. The real income of a family with four children is only about one-third higher than that of a childless couple, probably corresponding to a living standard about one-third lower. Secondly, the equalising effect of social policy appears to be much the same at different levels of income, although this conclusion should not be generalised for incomes outside the limits of the two examples taken in the tables. With higher incomes, a similar effect on total disposable income might still occur because of the increasing value of income tax concessions. As income rises, however, the value of the flat-rate benefits of the National Health Service and the Welfare Foods and school milk schemes would be of relatively less importance. In these higher income ranges, then, we might

expect families with children to fare worse relatively to childless households in the same income bracket than do similar sized families with a rather smaller income. Nor can the results be generalised for all family sizes. In the higher income groups, the family will fare relatively well until the number of children is such as to take it out of the income-tax range. At £650 a year, the family with four children pays no income-tax, so the arrival of a fifth child would benefit the family only to the extent of the flat-rate Family Allowance.

IV.

The idea has grown up in Britain that the State confers substantial benefits upon the members of certain classes, namely on those with relatively small incomes and on dependent families at most income levels. It is proposed here to examine carefully the nature and extent of these benefits and to compare them with the burden of taxation on representative households. In Table III, an attempt is made to assess the tax burden on the representative households already considered in the earlier tables.[1]

[1]The construction of Table III requires some explanation. Income tax liability is easily assessed. The figure for rates is based on H.F. Lydall, *op. cit.* p. 89, Table 179. (The figures relate to the year 1951-52, and therefore underestimate rate payments since no correction has been made). The major part of the National Insurance contribution has not been counted as a tax, since the benefits received under National Insurance have also been excluded from Social Income. The new entrant is, in fact, paying the actuarial value of the benefits received, and it is therefore reasonable to treat the Exchequer contribution to the National Insurance Fund as a payment towards the cost of giving pensions to earlier entrants beyond the actuarial value of the contributions made. The contribution from the National Insurance Fund to the cost of the National Health Service is treated as tax, however. Indirect taxation has been calculated by using the figures for expenditure of particular households in the Ministry of Labour's *Survey of Household Expenditure* (1954) and the ratio of the aggregate expenditure on various commodities to the tax collected on these commodities (Blue Book, *National Income and Expenditure*, 1959). The indirect taxes it was found practicable to allocate in this way were those on tobacco, alcohol, household durables and hydrocarbon oils (partially only in this last case). No allowance has been made for the fact that the expenditures recorded on alcohol and tobacco in the Ministry of Labour survey are known to be greatly underestimated. All these taxes, and some, like surtax, which would not normally be paid by any of the representative households considered amount to something like 88% of the total tax revenues of public authorities. It would not be easy, if possible at all, to make any reliable allocation of the remaining taxes. The remaining taxes are predominantly indirect, and since these may tend to be regressive, it is probable that by increasing the sum of the allocated taxes paid by a representative household in the ratio 100/88 to obtain a "probable total tax burden," that burden is in fact being underestimated.

TABLE III

TAXATION AND SOCIAL INCOME FOR REPRESENTATIVE HOUSEHOLDS

	Married couple with children as below :					
No. of children	0	2	4	0	2	4
Wage per annum	£450	£450	£450	£650	£650	£650
Indirect Taxes on :						
Alcohol	£9	£6	£7	£10	£8	£6
Tobacco	£27	£27	£30	£34	£30	£29
Household Durables ...	£4	£6	£4	£12	£8	£3
Hydrocarbon Oils * ...	£3	£2	£2	£7	£4	£3
Direct Taxes :						
Income Tax	£16	—	—	£61	£9	—
N.H.S. Contribution ...	£6	£6	£6	£6	£6	£6
Rates	£12	£12	£12	£15	£15	£15
Total of above Taxes ...	£77	£59	£61	£155	£80	£62
Probable Total Tax Burden	£87	£67	£69	£176	£91	£70
Total Social Income † ...	£24	£78	£164	£24	£78	£164
Net Social Income ‡ ...	—£63	£11	£95	—£152	—£13	£94

* It has been possible to allocate this tax only in so far as it is paid by the private motorist or the user of public road transport ; no allocation of the duty paid by hauliers, etc. and entering the costs of all goods has been attempted.

† As in Table II.

‡ The difference between Total Social Income and Probable Total Tax Burden.

It is clear from Table III that no net social income accrues to childless couples in a quite low income bracket, and that the net benefit for the two-child family is quite small, if not negative, at most income levels. Indeed, it is possible in view of the known under-reporting of consumption of tobacco and alcohol that even the two-child family with a wage as low as £9 a week is paying more in taxation than the value of the social services it can be sure of benefiting from.[1] This finding is in keeping with the studies of income redistribution that have been made by such authorities as Findley Weaver, de Jouvenal and Colin

[1]It may also be noted that with a four-child family, a man receiving a wage or salary of £650 a year may be paying roughly the same in taxes as a man with £450 a year and the same family.

Clark. The main difference from these other studies lies in the stress placed here on the effect of social policy on representative households rather than on broad social classes.

There is, therefore, ample grounds for the contention that there is comparatively little redistribution of income brought about through the fiscal system, and that what there is may be as much horizontal redistribution as vertical. This argument has been contested by Professor Peacock and Mr. Browning, but they are not so much contesting the validity of the statistical data on which it is based as suggesting that certain classes are in fact benefiting substantially from current policy in so far as they are in large measure relieved of any obligation to contribute to the cost of indivisible services provided by the State.[1]

It is clear that the standard of living of the dependent family is very much lower than that of the childless household in the same social class. Many might argue that this is only natural and that parents should not expect childless members of the community to share their burdens. On the other hand, there is this widespread belief that the State is already doing a great deal to help dependent families, and some resentment on the part of childless couples and single persons at the heavy burden of taxation they have to bear in order to provide Family Allowances and other benefits for dependent families. Table III shows how mistaken this belief is, and that little or no net assistance is given to families except where there are perhaps three or more children, a family size which has ceased to be typical. For smaller families, the most that can be said is that these have been relieved of part or whole of any contribution they might otherwise have been called upon to make towards the cost of indivisible government services. Unless one rejects the whole concept of taxation according to ability to pay, or artificially restricts this concept to income unrelated to family responsibilities, this is obviously desirable.

Moreover, the position of dependent families has worsened since the end of the war, both as a result of general economic conditions and changes in tax policy. Thus the index of retail

[1] "The Social Services in Great Britain and the Redistribution of Income," in *Income Redistribution and Social Policy* (Ed. Peacock), London, 1954.

prices rose from 100 in June, 1947, to 162·1 in 1957 (annual average). A special cost of living index for the three-child family with an income in the range of £10 to £14 a week would have shown a rise to 166·1 (or 167·5 in the case of such a sized family with an income between £8 and £10 a week). This reflects the fact that the lower income groups and larger families spend a larger proportion of their incomes on food, which has shown the greatest price increase over this period. This relative movement of the cost of living against the lower income groups and larger families has been a fairly consistent feature of the post-war period and cannot be attributed exclusively or even mainly to such changes in government policy as the abolition of food subsidies.[1]

Changes in taxation have normally given greater benefits to the higher income groups and smaller families. Thus the income tax reductions made in 1946 gave the childless couple with £500 a year a gain of £36 10*s*. whilst a couple with three children gained only £27 6*s*. In 1952, a childless couple with the same income gained £18 2*s*. from tax reductions whilst a couple with two children gained only £11 4*s*. The three-child family gained nothing since it had passed out of the range of income tax.[1]

V.

Whilst recognising that "social income" did not at present equalise living standards between different types of households, Lady Williams argued [2] that it could be made to do so. How far such equalisation of living standards should be carried is necessarily a value judgment. In certain professions and occupations it is a live issue. For many years, the difference in living standards between women teachers and their male colleagues with families has been a source of discontent to a vociferous element within the profession, though rather inconsistently this element has not regarded the difference in living standards between male teachers with and without families as being in any way anomalous. If leakages to the

[1] These details are taken from the paper by Mr. Hutchings already referred to.

[2] *loc. cit.*

press are any guide, a system of family allowances as part of the salary structure has been considered more than once by the Burnham Committee since the war.[1]

If, however, it is accepted that more ought to be done to help dependent families, the means to this end can be discussed without making too many value judgments. Such help for the dependent family would almost certainly be in the form of financial aid, for it is difficult to see how help in kind could be substantially increased without intolerable interference with the freedom of consumer choice. Apart from increased tax reliefs where appropriate, the alternatives would be increased Family Allowances on the present flat-rate basis and increased allowances which were related to income. So long as Family Allowances are financed from general taxation, there is probably a strong case for a flat-rate. It would be extremely difficult to justify the obvious differentiation involved in paying income-related allowances out of public funds. If, on the other hand, a system of family allowances were to be incorporated in the wage and salary structure, or financed by a special contributory scheme where contributions were also linked with income, the difficulties would be far less. The choice between the two systems would depend on the view taken of the purpose of family allowances. If the primary purpose is to ensure that every dependent family is able to enjoy a certain minimum standard of living (and this *was* the motive of the English pioneers of family allowances) a flat-rate system is what is needed. If, on the other hand, the aim is to raise the living standard of the dependent family relatively to childless households in the same social class or to prevent households suffering too large a fall in their standard of living with the arrival of children, a system of allowances linked to income would appear to be indicated.

In France, family allowances have been on a far more generous scale than in Britain, though still at a flat-rate. In 1949, the allowances paid were 3,050 francs a month for the second and 4,800 francs a month for the third and subsequent

[1]From April, 1961, women teachers in England and Wales will be paid at the same rates as men, but the 1959 report of the Burnham Committee made no reference to family allowances.

children. There was a further allowance of 2,400 francs for the first and second children and 1,200 francs a month for third child if only one parent were earning more than 4,000 francs a month. These allowances were financed by a 16% payroll tax on earnings up to 38,000 francs a month, a ceiling so placed that few *wage*-earners would exceed it.

The importance of the allowances can be seen by comparing them with this ceiling. Thus a family with two children and only the father working would have received 7,850 francs a month in family allowances or roughly 20% of the ceiling wage, and substantially more than 20% of the wage of the average manual worker. If account is also taken of tax reliefs, a family with two dependent children where the father was earning 175,000 francs a year would have had a disposable income of 269,250 francs. It would have had a disposable income 58% greater than that of the single person. In a comparison between a similar family and a single person where the wage or salary was 1,000,000 francs a year, the difference would have been only 24%.[1]

Under this system, where flat-rate benefits are financed by a proportionate payroll tax, families of lower paid workers are better able to maintain their standard of living relatively to childless households in the same class than are the families of higher paid workers. Before one advocates a system of this kind, one ought to consider its effect on the wage and salary structure. The relative overall attractiveness of a skilled job compared with an unskilled one would be reduced for the average young man who expected in the normal course of events to marry and have a family. The French system does not appear to have been analysed in this light. There are a number of questions that could be asked. Might the effect of family allowances in reducing the attractiveness of the skilled job have been offset by a widening of the differences between basic rates? Or has the reduced attractiveness of the skilled job had little effect on recruitment, perhaps because the

[1]Details are taken from G. Rottier and J. F. Albert, "The Social Services and Income Redistribution in France," in *Income Redistribution and Social Policy*.

earnings of skilled workers had included a substantial element of quasi-rent or monopoly revenue?

Any attempt to introduce such a system in Britain would involve difficulties. The financing of a scheme by a tax on employers might be regarded by them as unduly onerous. Financing it by a tax on the workers themselves might equally meet objections, especially from older men who would not benefit immediately nor stand to benefit in the future. Certainly, *any* attempt to improve significantly the lot of the dependent family in Britain to-day would meet considerable opposition. Nevertheless, it is important that discussion of policy should be based on a true understanding of the facts, and that popular misconceptions about the extent of present help to the family should be dispelled.

J. M. Jackson

University of St. Andrews,
Queen's College, Dundee.

Monopoly and Competition in Television: Some Policy Issues[1]

The new debate upon public policy toward television has centred on the award of broadcasting rights for a third channel. Should it go to the B.B.C., to a new public corporation, to a new or existing commercial organisation, or to toll television? More is at stake than the nature of the new channel since the additional service may fundamentally alter the behaviour and performance of the existing services and thus affect the pattern of development of the whole broadcasting system. This is one of those rare moments when the public has a real opportunity to choose among alternatives: the purpose of this paper is to explore them.

The procedure used is (i) to define the standards of judgment employed, (ii) to relate performance with respect to these criteria to the form of organisation by considering two prototypes of organisation, (iii) to sketch the possibilities of an ideal mixed system, and (iv) to compare such a mixed system with the existing British mixed system. In this context the uses of a third channel are then discussed.

Among the many standards of judgment that have been suggested the following will be used in this paper :[2]

1. *Provision of an ample and unbiassed source of news and an open forum for commentary and public discussion.*

2. *Existence of a sufficient number of independent sources of employment* for specialized personnel to prevent capricious or discriminatory treatment, and to provide a fair market for valuation of their services.

[1]I acknowledge with gratitude the helpful comments of several members of the staff of The London School of Economics, and the financial support of the John Simon Guggenheim Memorial Foundation and the Social Science Research Council.

[2]The omission from this list of the criterion "provision of a technically adequate national service" is justified because of the ample international evidence that it is consistent with such a wide range of alternative forms of market organisation as to be irrelevant to the subsequent discussion. The striking common characteristic of the very different broadcasting systems of Britain, the United States, Canada, Australia and others is their success in overcoming the technical problems of distance and interference.

3. *A commitment to public service.* This phrase is intended to embrace : attention to social objectives such as public information and education, stimulation and encouragement of creative activity, preservation of the distinctive national and regional cultural patterns, and observation of canons of good taste in entertainment and argument. For brevity it will subsequently be called *public service responsibility.*

4. *Provision of a rich and varied choice of programmes with due regard to the preferences of the population.* A wide choice of programmes at any time, and over time, is important for two reasons. The first is that tastes vary widely. The product of the industry—the programme—is one for which there is no inherent, unambiguous index of quality. Rational and informed consumers choose differently and there is neither need for nor likelihood of a consensus of judgment. Secondly choice is essential for the formation of tastes since it permits the experimental sampling of the unfamiliar. Given only two or three channels, choice is necessarily limited : the problem is to avoid further, unnecessary restrictions.

Each of these objectives is critically influenced by the structure of the industry and the motives of its members. Each is thus influenced by the question of public or private ownership, by the method of financing and by the extent and form of competition. Since there are several standards of judgment employed there is a real possibility of a conflict of criteria : a form of organisation that performs well when judged by one standard may do poorly by another. Indeed, this sort of conflict is the crux of the debate between the advocates of public monopoly and private competition in broadcasting. This debate is worth exploring, not to resolve it,[1] but to find if there exists a mixed system that achieves the benefits of both. For this purpose it is helpful to use as prototypes (*a*) a monopoly under a public corporation as in British sound broadcasting, and (*b*) a private, commercial, competitive system under overall public surveillance as in American television.

[1]It is impossible scientifically to resolve it without an agreement on the ordering (or weighting) of the objectives. Since I believe this agreement to be impossible, I believe it is unwise to force an election (by majority vote, or in any other way) if it is unnecessary.

While the pattern of operation of the B.B.C. in sound requires no description here, the following brief digression on the prototype American system may be helpful.

THE CHARACTER OF AMERICAN TELEVISION.[1]

Three national network organisations (N.B.C., C.B.S., A.B.C.) dominate American television. Each offers its affiliated stations about twice as many hours of programmes per week as B.B.C. television. Some of these programmes are sponsored, some not; of the former some are network produced, some produced by advertising agencies or independent producers. While the networks own a few key stations, most stations are independent, private, commercial organisations that have voluntarily affiliated with the networks, yielding to the networks options on a significant portion of their time, the major share of the revenue from the sale of their time to network advertisers, and much of their jurisdiction over programming. In return they are provided with the cable interconnection that makes simultaneous national broadcasting possible, with a stream of programmes of all types, and with the minor share of the revenue from the sale of their time as part of the network. The two biggest networks are enormously profitable by any standard, A.B.C. is less profitable, but is above the average of American corporations of its size. Stations typically enjoy moderate profits and are clearly better off with network affiliation than without it.

Stations, rather than networks, are licensed by the Federal Communications Commission and renewal, while subject to standards of overall performance, is virtually automatic. Stations receive additional revenue from spot commercial announcements and from sale of their time for local or non-network "national spot" broadcasting.[2] Non-network sources account for perhaps a quarter of the total time of stations and

[1] For more detail, see F.C.C. *Report on Chain Broadcasting* (1941); *Network Broadcasting* (1958) House Report 1297, 85th Congress, 2nd Session; P. O. Steiner, *Workable Competition in Radio Broadcasting* (1949) Harvard University Library.

[2] "National Spot" is the phrase used to describe the practice of purchasing a film series and telecasting it in a series of selected cities at varying times. Sponsorship may be by one national advertiser or by local or regional advertisers who co-operate in the venture.

more than half of their revenue. These fractions should not lead one to under estimate the importance of network programming for two reasons : firstly, virtually all of the high budget programmes, entertainment and public service alike, are network and their drawing power creates the audience at which both spot commercial announcements and the adjacent programmes aim ; [1] secondly, with only two or three stations in the typical market area, virtually all stations are affiliated to one of the networks and rely heavily on their network programmes for the "quality" programming with which they justify themselves to the F.C.C.

The networks' dominance, influence and affluence reflect a substantial degree of monopoly power *vis-à-vis* all other elements in the industry. They are insulated (mainly by past allocations decisions) from new networks or from other challenges to their dominant position. This position permits them to pursue a multiplicity of goals and gives them wide scope for decision. They do, however, compete actively with each other for the related (but not identical) goals of audience, revenue, and public prestige. This is the competition of rivalry —of the "big few"—rather than the impersonal competition of classical economics. External influences, such as the residual responsibility of stations, preferences of listeners and advertisers, and public opinion limit but do not wholly determine network practices. Since the limits are wide it is more of a mistake to be preoccupied with them than to ignore them. In the main, networks can do what they want. When they choose, for example, not to resist advertiser pressure it reflects lack of courage or desire rather than lack of power.

Competition, Monopoly, and the Public Interest : Prototypes Compared.

1. *An ample and unbiassed source of news, and an open forum for commentary and public discussion.*

The arguments against monopoly power in these areas need no rehearsing here. Nor is it the fear of outrageous abuse of

[1] The value of "adjancies" (as they are called) to a station from a particular popular network programme may exceed its direct revenue from it.

freedom of thought or expression that motivates opposition to a benevolent and public spirited monopoly like that of the B.B.C. in sound broadcasting. It is, rather, that the centralization of authority over the selection and presentation of public issues imposes an awesome responsibility for a benevolent monopoly : it must be prudent and it must be studiously impartial. This responsibility leads, apparently inevitably, to paternalism. At best, and the B.B.C. frequently achieves this best, it produces the broadcasting equivalent of a superbly balanced book. But this is not the equivalent of a balanced library. The brilliance of individual items may obscure the fact that there is no appeal from the decision *not to cover* a particular story, or to give it merely routine coverage. This can of course occur under any system, but is less likely if there are several independent decision units than if there is one.[1] This is a major objection to monopoly in broadcasting.

How does American experience compare? With respect to news reporting it is by usage free from advertisers preview or editing. Both unsponsored and sponsored broadcasts are common, at both the station and network level ; selection and production decisions reside with the broadcasters in all cases. The frequency of such news broadcasts is great, and because of the diversity of the selection, the number of stories reported in a given day is substantially larger than on the B.B.C. News commentary is also frequent and exhibits substantial range of opinion—from the extreme right to the middle left.[2] While puditry far outweighs polemic, both are present. In the field of public discussion, panel interviews and documentary coverage

[1]A recent American experience is instructive. During Kruschev's attendance in New York during the fall of 1960, a major boycott on interviewing him was agreed to by the networks, on a suggestion from Washington. This boycott was broken by an independent station in New York and the resulting broadcast attracted international attention. At stake is not the wisdom of the particular decision, but the character of an open society. Compare the decision of the B.B.C. not to broadcast the results of the American presidential election throughout election night with its decision to broadcast the heavyweight fight between Patterson and Johansson at 3-15 a.m. for an example of the capriciousness of a monopoly.

[2]Most broadcasters and commentators are regular staff members of the station or network news departments. A few commentators are chosen by the sponsors which include trade unions as well as more traditionally conservative organizations.

in depth are frequent ; once in a while in bad taste, typically informative and occasionally superb.

It is necessary to note that while pluralism of decision making inheres in competition, they are not the same thing, and it is the former that is more important here. Limited competition appears to be more satisfactory than atomistic competition : the two largest networks are the clear leaders in quality news and discussion programming. This is partly due to their intense competition between themselves for public prestige and partly to their secure (monopoloid) financial position which permits maintenance of an expensive public service establishment. A.B.C., the weakest of the networks, falls well short of N.B.C. and C.B.S., and the multitude of individual stations lag behind the networks. High quality coverage probably owes more to affluence than to competition, but they are not wholly incompatible. And even limited competition avoids the major objections to monopoly in this area.

2. *Existance of a sufficient number of independent sources of employment.*

The market for television performers, directors, producers and technical personnel is necessarily discontinuous and imperfect. The advantages of having even two or three potential employers rather than one are obvious. They are particularly compelling where artistic judgments are involved. Blacklisting, favoritism, and discrimination may occur, but they may also be imagined if there is no real alternative source of employment. Pluralism of decision making provides an essential element of protection.

Because skills are specialized and market are imperfect, appropriate levels of compensation are difficult to determine. It is clear that introduction of independent television in Britain has led to a significant increase in the pay scales of the B.B.C. as a response to the loss of personnel to its competitor. This suggests that monopoly, public or private, exercises significant monopsony power as well.

If one accepts the view that these considerations are important, it is clear that competition—even competition

limited to a few competitors—has a major advantage over monopoly.

3. *Public service responsibility.*

A semi-public monopoly has one enormous asset : it is free to undertake a wholehearted commitment to public service. Freed from any crass scratching for revenue, freed from competition for listeners, and insulated from pressure groups, it may consciously and conscientiously pursue its image of the public interest. The B.B.C's world-wide reputation and its many programming triumphs attest to the public benefits that can be achieved.

These benefits are not without costs. The image served is that of the monopoly. That it may be paternalistic, that it may identify too closely with the "establishment," that it may resist changes in itself and become impervious to changes in the public climate, are real rather than phantom dangers. They are the hazards of power wherever it is found. Complete lack of power, the characteristic of atomistic competition, while avoiding the dangers, negates the benefits as well in cases where public service is not identical with private interest. The challenge is to find the mixture of monopoly and competition that gives the best balance.

The American mixture does not. Not that the conscious pursuit of public service is absent: the networks, do have such an image and it is reflected in news and documentary programmes, and elsewhere as well. But even networks can afford to be "uneconomic" in some areas only if they attend to the business of collecting revenue most of the time. If one makes a rough distinction between entertainment and other programming it is only slightly unfair to say that networks abdicate public service responsibility in the former and embrace it in the latter. As to entertainment, the creed is to "give them (the listeners and/or the advertisers) what they want." (How well they do this is discussed below). For the other they can be magnificent. As examples consider their decisive role in the destruction of McCarthy by the simple expedient of relentless coverage of his investigation of the Army ; the coverage of political campaigns and elections ; the C.B.S. documentary

on Polaris; the N.B.C. documentary on the U-2, with its restrained if devastating exposure of the ineptness of the government then in power. These are the equals of the best of the B.B.C.

But these activities are inversely related to competition. A.B.C. has gained both listeners and profits by leaving prestige programming to its rivals and developing "adult" westerns. National spot programming eschews it entirely, and individual stations frequently reject unsponsored network prestige programmes in favour of national spot or locally sponsored film series if they have an option. Subjecting the networks to fiercer and more compelling competition, a goal frequently advocated and fruitlessly pursued by the F.C.C., would work against rather than for public service responsibility.

The successes of the networks should not detract from the question of whether enough of the networks' time is devoted to programmes of this sort. This is the problem of the programme mix.

4. *The provision of a rich and varied programme.*

Here the disadvantages of limited competition are most pronounced. This fact, and the reasons for it are very important but only dimly appreciated. It is convenient to consider the problem in stages.

Consider first a situation in which there are three channels operating concurrently for a single period (say 30 minutes) broadcasting to a population of potential listeners whose tastes vary in the following very simple way: out of every 100 potential listeners 80 want a programme of type A (e.g. "western"), 18 want a programme of type B (e.g. dramatic show), and the remaining 2 want a programme of type C (e.g. discussion of current events). Further assume to start with that a listener will turn on his set only if a programme of the type he wants is broadcast.

A monopoly interested in maximizing the number of listeners (which in this case is the same as satisfying the largest number of people) will clearly produce one programme of each type, one on each channel, and both capture and satisfy the whole audience. Three competing firms, each trying to get as

many listeners *for itself* as it can, will probably *all* produce a programme of type A for the simple reason that each one's share of the mass audience will be larger than the whole of the potential audience for programmes of types B or C. If we assume that two or more stations producing the same programme type will share the audience equally, it is easily seen that a fourth channel will also produce type A (20 per cent. of the audience) rather than type B (18 per cent.). Indeed, one would need 5 stations before one could expect type B to be produced, and at least 48 channels before type C would appear attractive. Duplication is a consequence of competition for listeners, and if choices are varied and the number of channels limited, duplication will leave some listeners unnecessarily dissatisfied. While each of the duplicating programme producers can claim that "we give people what they want," this is not so in the aggregate.

If viewers will accept (if grudgingly) less-preferred programmes rather than none at all, the tendencies for duplication under competition are increased. To see this quickly, suppose 70% of the audience prefer type A and 30% type B, but will listen to either. If there are only two channels, both will produce type A and each will get 50% of the audience. A third channel has a choice of producing type B and getting 30% of the listeners, or duplicating type A and getting one-third of the total. If it aims strictly at maximizing its audience it will choose type A. (If viewers had not been shiftable, it would have chosen type B).

When attention is directed to variation over time rather than to a single period, there are clearly substantial opportunities for increased variety. But there are major forces tending toward duplication as well, and a clear overrepresentation of the most popular programme types. This can be shown theoretically, but the argument is somewhat complex, even tedious, and it will be omitted.[1] Empirically, duplication of programme types by channels at the same time, and also over time, are perhaps the most clearly visible features

[1]For extended theoretical treatment of this and other complications, see P. O. Steiner, "Programme Patterns and Preferences and the Workability of Competition in Radio Broadcasting," *Quarterly Journal of Economics*, Vol. LXVI (May, 1952), pp. 194-223.

of the American system. The near-saturation of prime evening hours on all networks with westerns, "private eye" shows, and situation comedies has been noted by many critics. The cost is less coverage of minority tastes than their place in the preference pattern of the population warrants.[1]

So far, then, as satisfying as many listeners as is possible is a sensible objective, and so far as choice is only between programme types, it is clear that complete monopoly is, in principle, a more effective instrument than maximum competition when there are binding limits on the number of competitors possible. Monopoly, even private monopoly, is motivated to avoid duplication ; competitors are not.

Some will quarrel with the objective. Broadcasting, they will argue, should strive virtually exclusively to edify and to educate rather than to cater to the wants of its audience. Whether or not one shares this view (I do not) it is clear that monopoly is again much the most effective way to achieve it, although a monopoly need not act this way. To carry this view to an extreme, the case can be made for only a *single* channel. While there are times (such as a major address by the head of state) when a case can be made for silencing distractions, this practice will find little general support in the basic tenets of an open society.

Viewers select specific programmes as well as programme types. A choice between two westerns (or two religious programes, or two variety shows) is, after all, better than no choice. Not only do tastes vary within programme types but competition between programmes imposes standards of competence on programme producers. But such choice is expensive if it precludes the wider alternatives. Further, the same forces that tend toward duplication of programme type lead to imitative programming where competition for listeners is the spur.

[1]Nor does this observed behaviour fully reflect the working of the simple assumptions of the above examples. Networks do temper their uninhibited search for listeners because of their sense of public service and their desire for critical esteem, which their partial monopoly allows them to indulge. Comparison of network programming with syndicated and local programming provides ample support for the networks' repeated assertion that they devote more time and money to providing both variety and quality than mere profit-seeking would dictate.

The performance of our prototypes with respect to the quality of individual programmes is not easy to assess, because of the presence of differences in the form and adequacy of financing available. The American prototype taps additional sources of revenue and is much more generously financed. The results of this show, not only in the higher budgets available for popular entertainment but elsewhere as well. For example the extensive (and in my view splendid) coverage on a regular basis of top sporting events, professional and amateur, occurs because the system generates revenues sufficient to compensate the promoters for the declines in attendance. The price paid for this affluence is the foothold given advertisers in choice and character of programming. The results are some of the best and worst programmes in the world. It is not necessary to ask British viewers to accept this assertion on faith : they see a generous sampling of American television on their sets, somewhat biassed in favour of the mediocre.

In summary the comparison between the prototypes—public monopoly and private, commercial, limited competition —shows :

(i) the advantages of pluralism in providing an open forum for news and discussion and in providing alternative sources of employment give competition a decisive edge.

(ii) Public service responsibility flows from two sources : power and motivation. The public monopoly has both, and has the advantage. But the form of limited competition considered leaves substantial monopolistic power in the hands of the competitors who can (and do) choose to compete for public esteem as well as for audience. In either case the existence of a protected position provides the opportunity for conscious public service.

(iii) Monopoly is inherently motivated to provide the complementary programming that leads to maximum variety ; competitors are not so motivated and a significant restriction on choice owing to both duplication and imitation is to be expected.

(iv) Competition, if it generates substantial additional resources for programming, can do some things better than a (relatively)

under-financed public monopoly, and (while it should not be overstressed) the need to compete for audiences may have a beneficial effect on the particular programmes of the kinds most heavily demanded.

If choice is limited to one or the other of these prototypes, judgment as to a "balance of advantage" is required. That there is no need so to restrict choice is indicated by the present existence of a mixed system in British television, and it is perfectly safe to predict that it will remain mixed. The challenge is to find a mixture that can tap the advantages of both.

An "Ideal" 2-Channel System.

Suppose that one channel is given to an adequately financed public corporation, called the Public Broadcasting Co. (P.B.C.) and that the other is licensed or leased to the Commercial Television Corporation (C.T.V.), a private company, organized for profit, that sells advertising as its chief source of revenue.

Subject to such overall terms as the appropriate authorities may designate the C.T.V. is encouraged to pursue its own interest, which we may crudely designate as seeking to maximize its audience. It will, in the main, cater to the dominant preferences of its audience, although if care is taken in granting of the licence there is no reason to suppose that the drives for prestige and status will be entirely absent. The form of commercial activity permitted (sponsored programmes or commercial announcements) is left to the licensing authority; if adequate financing is possible without sponsorship, it seems preferable.

The P.B.C. is instructed to conduct itself so that the *total* broadcasting service (C.T.V. *and* P.B.C.) is as rich and varied as possible. During hours in which both channels are operating this will mean that the P.B.C. will complement rather than compete with C.T.V. and provide the best possible *choice* for those not attracted to the C.T.V. If the private channel produces or imports splendid westerns or variety shows, the P.B.C. will provide opposite them a real choice, hour by hour. Of course, where the P.B.C. feels the C.T.V. is doing a poor job it is free

to compete : to discipline, if you like, the commercial channel and force the absolute standard of its programming up to a higher level. What the P.B.C. will *not* do is to schedule a Sunday night drama series opposite a similar series on C.T.V. (although it may do so at another time, opposite a variety show, if it thinks wise). Nor will it cover horse racing at Newbury while C.T.V. is covering racing at Windsor. During hours when C.T.V. is not broadcasting—P.B.C. is licensed to broadcast from early morning to midnight—it will act as the public monopoly that it is, with due regard for the nature, needs, and desires of its potential audience.

While the P.B.C. will thus be imbued with its role as a complementary service it will not always aim for the second largest audience—it will vary its programming to appeal, at least sometimes, to all minority preferences, and will not always present the same choices. Further it will be quite free to over represent minority tastes if it believes that public information, education, or welfare requires it. In short, it will do very much the same thing as if it were the second service of a *public monopoly* in which the programme of the first service were previously settled by higher authority. Co-ordination does not require mutual agreement ; it merely requires that one participant be adaptable and properly motivated.

Two consequences, at first glance unfortunate, will emerge almost by definition if P.B.C. plays its role properly. First it will surely lose the battle for audience ratings most of the time (there may be occasional surprises) since it is inviting C.T.V. to cater to the largest audience. This should neither discourage nor deter it. Its job (as a crude first approximation) is to maximize the total audience, not its *share* of the audience. The danger here is that the financing authorities will take this as a measure of failure and withhold funds. Perhaps ideally, the P.B.C. could have its revenue geared to the *total* satisfied audience, not the potential audience, and thus be motivated to play the desired role. The second consequence is that neither the C.T.V. nor the P.B.C. will provide a fully balanced diet. The former will be biased toward popular entertainment, the latter away from it. Some may object to this lack of balance in the offerings of each channel-authority, but, surely, this

would miss the point. Viewers are not required to choose a *channel* for the duration, but to choose programmes. Comparative shopping is easy and cheap (it would be easier and cheaper if the P.B.C's publication *Broadcasting Times* provided information about programmes on *both* channels). Choice between channels, instead of between programmes, greatly reduces options, and it should be discouraged rather than fostered by the public-serving P.B.C. In this connection the P.B.C. will take pains to avoid unnecessary impediments to choice owing to scheduling times: if the C.T.V. schedules programmes to begin on the hours and half hours, P.B.C. does the same, rather than on the odd quarters.[1]

In what senses is this "ideal"? First of all it provides for the complementarity of programming that is a major advantage of monopoly over competition without forfeiting competitive prodding when deemed necessary. Secondly it retains the unique organization dedicated to public service that is possible only if financing is at least at one remove from the pressures of commercialism. Thirdly it provides the pluralism that is absent in monopoly. And finally it provides an ample source of non-public financing and uses it where it can do the job well, reserving the public funds for those things that private enterprise can not do, is not motivated to do, or does indifferently. In short it achieves the advantages of complementarity and public service responsibility without loss of the advantages of competition and of private financing.

The Mixture To-day.[2]

Does British television, in 1961, approximate this hypothetical system? It does not, as every viewer knows, and

[1]This is not intended to discourage programmes of unconventional length. An appropriate regulation might be that all programmes of thirty minutes or more *begin* on the hours or half hours.

[2]In this discussion, I group the various organizations involved in independent television under the single category "I.T.V.", notwithstanding the different organizations involved (among them, Associated Rediffusion, Granada, A.T.V., A.B.C., I.T.N., and the Independent Television Authority). This is not because there is no difference between the actual system and one in which all independent television is in the hands of a single commercial organization, but because the differences are not critically important for the central problems at issue. One major advantage of the decentralized system is the wider choice of employers it offers.

as the programme time-tables reveal. I.T.V. is like C.T.V., but B.B.C. is not P.B.C.

Consider Sunday, 1st January, in London (for which a diagram of programming is attached). This was not an average day—not even an average Sunday—because of the special holiday programming, well publicized in advance. Notwithstanding the exceptionally large number of hours during which at least one of the channels was operating (13), and during which both were simultaneously operating (9), the real choices available were severely restricted by duplication and by impediments of scheduling. With reference to duplication, notice, among others, the religious programming around the noon hour, and from 7 to 7-25 p.m., the feature films on both channels in mid-afternoon, and the original plays on both channels in the evening. With respect to impediments, note, for example, that the viewer choosing to see the Phil Silvers Show on B.B.C. (25 minutes) was forced to miss the last 10 minutes of Free Speech, and the first quarter-of-an-hour of a film in which build-up is critical to the plot, effectively, 100 minutes of programming was impeded. Or, again, how children glued to Pathfinders to Mars were forced to miss the start of Beauty and the Beast, surely designed for family entertainment.

All is not lost, of course. Viewers had a relatively clear choice of Concert *v.* Hawkeye in the late afternoon, of Showtime *v.* Danger Man, of What's My Line *v.* Val Parnell and of Monitor *v.* Maverick during the evening. Out of roughly 13 hours during which programming was available, there was no choice at all for 4 hours, choice between basically similar programmes for another 4 hours (of which about half were seriously impeded) and substantial choice for 5 hours (3 hours partly impeded). This on a gala day for telecasting! Surely it is possible to do better. Readers may enjoy attempting to rearrange the B.B.C. schedule to try and maximize the choice available, given I.T.V. programmes.

SUNDAY, 1st JANUARY, 1961 (London)

Time	B.B.C.	Time	I.T.V.
11-30 a.m.		11-00 a.m.	Religion (Service of Dedication)
12-40 p.m.	Religion (High Mass)	12-05 p.m.	
1-15	Farming		
2-00	Welsh Language	1-45	Discussion (Free Speech)
2-25	Comedy Series (Phil Silvers)	2-10	Film (Face In the Night)
4-05	Film (Two Tickets to Broadway)	3-25	Play (Interplay : The Talker)
4-50	Concert	4-15	Western Series (Hawkeye) (re-broadcast)
5-00	Childrens : Lennies Den	4-45	Children's Serial (Pathfinders to Mars)
5-40	Play (Beauty and the Beast : Part 1)	5-15	Travel Series : (Introd. to a Journey)
6-00	Misc. (In the Skies of Paris)	5-30	Advertising Magazine
6-10	Misc. (Sunday Special)	5-50	Misc. (Birthday Honours)
6-15	News	6-05	News
7-00		6-15	Misc. : Sunday Break (Documentary about Medical Missionaries, Discussion, Cartoon film)
7-25	Religion (Meeting Point)	7-00	Religion (Bands of Love)
7-30	News	7-25	News
8-15	Variety (Showtime)	7-30	Adventure-Spy (Danger Man)
8-45	Panel Game (What's My Line)	8-00	Variety (Val Parnell)
10-00	Original Play (Sunday Night Play : Nearer to Heaven)	9-00	News
10-10	News	9-05	Original Play (Armchair Theatre : A Headful of Crocodiles)
10-55	Interview : (Monitor)	10-10	Western (Maverick)
11-05	Epilogue and Weather	11-00	Current Events (Crystal Ball, 1961)

Some summary statistics for the entire week of 1st January to 7th January are given in Tables I and II. They reinforce the impression that at present the two existing channels fall far short of effectively utilizing the opportunities for choice.[1] All signs point to the fact that the B.B.C. is choosing to behave like a competitor (albeit a non commercial one that is not insensitive to public service responsibility). It competes actively—even frantically—for *share* of the audience rather than trying to provide a total service that is as good as possible. It is much closer to an American network than to the imaginary P.B.C. Since I.T.V. is also like an American network (though with some differences) the two are fairly close together : while there are real differences between B.B.C. and I.T.V. they can easily be exaggerated. If there is disquiet about programming it is not because of the absence of exceptional programmes on each channel, but because the overall programme fare seems thinner, more repetitious and more predictable than necessary.

TABLE I

HOURS OF TELEVISION BROADCASTING, LONDON
(Week of January 1st–7th, 1961)

	Hours 7 days	Percentages		
		7 days	Sat. & Sun. 2 days	Mon.–Fri. 5 days
Hours Available (8 a.m.–12 m.) ...	112	100	100	100
No broadcasting	47	42	24	49
Welsh language only ...	2½	2	3	2
One English language channel only	13¼	12	13	11
Both channels	49¼	44	60	38

[1]The distinctions between "impeded" and "clear," and "same" or "different" are of course subject to judgment. To illustrate from the Sunday programmes, the whole period from 2-0–4-15 p.m. is classified as impeded, but during the period 8-0–9-0 p.m., the middle 30 minutes are regarded as offering unimpeded choice, since interruption of a variety programme is not intolerable. In fact much of what is classified as clear is partly impeded.

TABLE II

ANALYSIS OF CHOICES AVAILABLE IN LONDON DURING HOURS WHEN BOTH CHANNELS TELECASTING IN ENGLISH

(Week of January 1st–7th, 1961) (Percentages)

Type Programme	Ability to Select		Total
	"Clear"	"Impeded"	
"Same"	21	5	26
"Different"	43	31	74
Total	64	36	100 (= 49¼ hours)

It has been suggested that B.B.C. and I.T.V. are similar not because I.T.V. has pulled B.B.C. "down," but because B.B.C. has imposed standards of telecasting that force commercial television to raise its sights. This argument draws no support from a comparison of I.T.V. with the leading American networks. In my (subjective) view it is clearly inferior in its overall performance to at least two of them. Others may disagree but the case must be made rather than assumed.

Why has the B.B.C chosen the course of competition for listeners? Apparently it is the victim of two delusions: that its success is measured by its relative share of the audience, and that it can persuade viewers to choose allegiance to a network. If true, these would justify all the tricks of the trade to capture an audience which could then be exposed to its "high" level programming. In this view the "other" service is the enemy, to be officially ignored and vigorously fought if it is not possible to pretend it does not exist.

The B.B.C's (in my view) mistaken conception of its role is not entirely of its own making. It does appear to suffer from serious under-financing, despite its very short broadcasting week (less than 60 hours). Its programming triumphs show its ingenuity in low-budget programming. Yet it can tap no willing source of additional revenue but must rely on the public purse.[1] If the public is going to take audience share as the

[1]The license fee is, after all, a tax—the B.B.C's share is merely a semi-automatic appropriation.

measure of success, the B.B.C. must either conform or persuade the public to change its view. To do the latter it must first concede the fact that independent television has a role to play ; something it has so far failed to do.[1]

While Table II highlights the failures of the B.B.C. in complementary programming, Table I focuses on the limitation of choice owing to the partial utilization of the two channels. Sound broadcasting is offered for 100-120 hours per week ; television on *either* channel for 65 hours and on both channels for less than 50 hours. Millions are home by day—housewives ; the retired, disabled, and ill ; the unemployed ; children on holiday and pre-schoolers ; and off-shift workers. The argument that only peak listening hours matter is a form of responding only to the majority needs that is analogous to catering only to majority tastes. Data on sound broadcasting indicates the existence of a large off-peak audience ; whether or not it is large enough to justify commercial television, the case for public provision of a public service seems clear.[2]

Without any increase in the number of channels it would be possible greatly to increase the range of choice by increasing the funds available to the B.B.C., by increasing the hours in which it is permitted to telecast and—most important—by persuading the B.B.C. to alter its image toward that of the imaginary P.B.C.

Starting from scratch one might easily create a P.B.C. and perhaps intensify its motivation by financing it (in significant part) through a tax on the advertising revenues of the commercial channel. After a decade of operation of a mixed system this is not so easy. Not merely have habitual patterns of behaviour become instilled, but precedents, vested rights and articulate interests have developed with the tacit sanction of public policy. The real opportunity of the third channel is that

[1]One incident is revealing. After announcing in the *Radio Times* that it would carry exclusive films of an American boxing match involving a British fighter, the B.B.C. announced that unfortunately the arrangements had fallen through. It urged its disappointed viewers to switch to B.B.C. radio for a verbal description of the fight. What it neglected to mention was that the films were being shown on British television—on independent television.

[2]The American experience suggests that daytime commercial television is profitable at advertising rates well below evening rates.

it provides the vehicle for rectifying past mistakes and for achieving whatever pattern of television broadcasting is desired. This last clear chance lies with the Pilkington Committee ; the real cost of the wrong decision is the opportunities that are lost, probably permanently.[1]

THE EFFECTS OF A THIRD CHANNEL.

A third channel will, according to the way it is allocated, affect the motives and behaviour of the existing channels as well. In net effect it may create a system which, alternatively :

(i) Consists of three channels actively competing for share of audience, not unlike the American prototype, whatever the forms of financing ;

(ii) Has one truly complementary service, and achieves the basic benefits of the "ideal" 2—channel system ;

(iii) While not introducing a thoroughly complementary service, adds a new form of programming that does widen choice ;

(iv) Both creates a complementary service *and* adds a new *third* alternative type of choice.

While any one of these changes represents some improvement, in terms of public objectives, the first seems least attractive, and the fourth most promising. Between the second and third, opinions may differ since the benefits and costs are distributed differently.

The proposal to add a second commercial channel, like I.T.V., falls into the first, least promising, category. It would necessarily be competing for listeners with both existing

[1]Official enquiries into public policy toward an industry are necessarily infrequent. When they occur they condition and constrain policy in the future. There comes a time after which aspects of these policies become virtually irreversible. In the British sound broadcasting experience, the public monopoly policy, recommended by the Crawford Committee in 1925 was perhaps reversible as late as 1936 : after the Ullswater Report and the House of Commons debate in 1936 it was, for practical purposes a closed issue. The post-war debates (1946 and 1950) on the wisdom of this policy while critically influential on the emerging *television* industry, left the mature sound service, unchanged. A careful survey of the American regulatory experience shows that the last clear chance for major changes in radio and television occurred in the mid-thirties and 1948-52, respectively. Unfortunately the major investigation into radio broadcasting was not reported until 1941 and into television until 1958—too late to leave any real alternatives open to the policy makers.

channels, and for revenue with I.T.V. Importantly, it would do nothing to induce the B.B.C. to abandon its competitive policy. While further commercial competition will provide more programmes and more resources for programming, it will fail to utilize the opportunities of using the new channel to achieve the benefits of a mixed system.

No more promising is the proposal to create a second public corporation, like the B.B.C. but independent of it. Two public channels competing for public revenue would find it impossible to resist competing for listeners with each other and with the commercial channel as well. Additionally this proposal would require at least a doubling of the public contribution to telecasting or risk a serious dilution of the resources available for programming. Whether the benefits would justify the increased costs seems questionable.

The surest way to create one truly complementary service would be to charter a new public corporation with the orientation of the P.B.C. That is, introduce by addition the complementary service that B.B.C. ought to be, but is not. This leaves to the B.B.C. an avowedly competitive role, and would surely cause the case for continuing public financial support to be re-examined. The logic of public support of a service competitive with and similar to the commercial channel rests not in the effects on the programme mix but on an ideological dislike of private enterprise, an æsthetic dislike of commercials, or both. The cost of indulging these dislikes is the size of the required public contribution. To suggest supplanting the B.B.C. is only partly facetious ; if it will not voluntarily play the role of the public channel and if it can not be persuaded or cajoled into doing so, it should be supplanted as the residual guardian of the public interest. Could a new public channel be given the desired motivation ? With the force of official debate and recommendation, with a proper charter, and with a form of financing explicitly independent of share of audience it would not be too difficult.

The B.B.C's proposal that *it* be given the third channel goes part of the way toward achieving a complementary service. The essence of the suggestion is that, with two channels, B.B.C. could compete with I.T.V. with the right hand while it provided

a true second service with the left hand. Its merit is that at least two of the channels would be operated in a complementary manner. This is not the same as a public service complementary to the combination of the other two channels, for there is nothing that recognizes the substantial role of independent television in providing service to viewers. At best a two channel B.B.C. would ignore I.T.V. ; at worst it would systematically impede it in an attempt to divert audience to one of its programmes. But, at least, the grant of a second channel would induce the B.B.C. to do some of the things it has not chosen to do with only one channel. The drawbacks are, firstly, the cost of financing two services one of which is avowedly competitive ; secondly, the lost opportunity to have three substantial choices ; and, one might add thirdly, the distastefullness of awarding the third channel not as a reward for past virtue but as a bribe to overcome past poor performance.

If one use of the third channel is to try and approximate the results of an "ideal" two channel system, an alternative is to introduce a thoroughly new form of television. The most frequently discussed is toll television. The promise of toll television lies in : (i) its ability to tap new sources of programming, previously beyond reach because of extraordinarily high costs of production or because of the need to compensate other industries for loss of revenues caused by the grant of broadcasting rights ; (ii) its ability to finance itself from the willingness of consumers to pay ; and (iii) the impetus to quality provided by competition for the consumers support with other uses of consumers income.

A toll service will not by itself provide the incentives for the existing channels to act in a complementary way, nor is it without problems of its own. The principal danger is that it will skim the cream from existing programmes and programme producers rather than enrich the mixture. One suspects that viewers would be willing to pay something for much of what they now see free, and this willingness could easily be exploited by toll television.[1] This danger exists if toll television is an

[1]The expectation of such exploitation has been the major deterrent to the introduction of toll television in America. The fear seems to me justified on a reading of the testimony before the F.C.C.

independent service, through its ability to offer premium fees to particularly popular actors, playwrights and others; it exists equally if toll television is part of an existing service through diversion of the most productive effort into the profitable pay-off field.[1] Thus the form of organisation of a toll service is of real importance if the promise is to be achieved and the dangers avoided.

But even a model toll service does use up the third channel. It may well be that (considering practical alternatives only) the British public is faced with a choice between (1) using the third channel to enrich programming along the lines of toll broadcasting or (2) using the channel to create a genuine public service network. If this is the real choice, it should be made explicit.

Is there any possibility that toll television can *both* add the genuine third alternative and also reorient the existing channels along the lines of the "ideal" 2-channel system? I do not know, but the possibility deserves discussion.

In order to provoke this discussion, consider the following possibility :

1. Establish a Public Television Authority (P.T.A.) with overall jurisdiction over telecasting and with control over the third channel. Included in the Board of Directors would be representatives of the government, of B.B.C. and of I.T.A.

2. Among the responsibilities of the P.T.A. would be determination of the broadcasting hours for each of the three channels, and the formulation of such general regulations as a co-ordinated broadcasting system require.

3. The third channel would be leased for toll television for such programmes or programme series as the P.T.A. deemed appropriate. In order to minimize the risks of diversion, the

[1]One suspects that the possibility of higher fees influenced the decision of British Actors' Equity to urge granting the third channel to toll T.V. The American experience again bears, somewhat indirectly, on the diversion problem. The introduction of "Spectaculars" on the American networks a few years ago killed a number of the most popular programme series : the actors, and others, wanted the prestige (and spectacular fees) of the star performances.

P.T.A. might quite properly demand a showing that neither of the other channels was willing to take on a particular programme and meet its costs before granting toll broadcasting rights. The P.T.A. would collect a royalty on toll broadcasting designed to capture the bulk of the extraordinary profits assumed to inhere in toll television.[1] It might, in addition, receive the portion of the television set licence fee devoted to television broadcasting, and (if required) a tax on the net profits or revenues of the commercial licensees under I.T.A.

4. The revenues of P.T.A., from all sources, should give it a substantial surplus in its own operations and would be used to support public service telecasting in a variety of forms. Top priority would be general support of the B.B.C. which would be freed from the need, real or imagined, to compete for *share* of the audience with I.T.V. Support might also be given to an expanded I.T.N., to encourage news and documentary coverage within commercial television. Further grants might be made in support of particular programme series, and even to individual programme producers.[2] Finally, support would be given to educational television.[3]

Whether any such scheme would achieve the objectives of three genuine alternative sources of programming and of co-ordinated programming, is worth debating. And there are, of course, other possibilities. It should at the very least focus attention on the central fact that, since the different alternatives promise very different results, the public has a real and important choice to make. What is essential if the major opportunity is to be seized is to get the co-ordinating benefits of monopoly, the deep sense of public service responsibility of the public corporation, the rivalry and open access of competition, and the maximum range of programming possible. Americans have never had the chance to sample the possibilities of a mixed system. The injection of an American type network

[1]Royalty agreements would be subject to periodic re-negotiation.

[2]If such programmes were sold commercially, the grants would be recapturable.

[3]Since toll programming would tend to be concentrated in the evening hours and on weekends, the P.T.A. might make the facilities of the third channel available for free broadcasting to, and by, schools and universities.

into British telecasting in 1950 failed to produce the benefits of a mixed system. The formation of the Pilkington Committee and the possibility of the third channel provide one last, clear chance. Nothing could be more appropriate than for Britain, 25 years after the start of its television service, to pioneer once again.

PETER O. STEINER

University of Wisconsin.

An Analysis of Post-War Labour Disputes in the British Car-Manufacturing Firms

Introduction.[1]

The car-manufacturing industry is only one sector in which the general incidence of strikes has significantly increased, over the past few years, from the relatively low level which it maintained for two decades after 1932. But it is a sector in which that increase has been particularly marked. Indeed, in the motor industry the curve of industrial unrest anticipated the general rise, which did not become definite until the mid-fifties : but the incidence of disputes in the car manufacturing firms showed a sharp upward twist over 1951/52. And whereas the increase in industry generally has been largely due to the re-appearance of occasional big "official" stoppages of the type that had virtually disappeared during the previous generation, the actual frequency of strikes in the car firms has also risen at an accelerating pace—so that the annual number of disputes,

[1]Acknowledgements are due to the Ministry of Labour's Statistics Division and to certain press libraries for help in accumulating basic data—and to Miss Theo Cooper in the labour involved. Treatment here is largely statistical, but the valuable background information given by several connected with motor industry unions or management should also be recognised. No responsibility for the result rests on any of these helpers.

This paper is derived from work largely completed before the Minister of Labour convened, in February, 1961, meetings of employer and union representatives to discuss the motor industry's labour relations ; and it is written while those meetings are still in progress and without foreknowledge of their results. Press accounts suggest they have so far concentrated on what the Minister has called "fire prevention"—communication, consultation and the handling of disputes in the workplace : which, since most disputes start at that level, seems fair enough. At the same time, improvement in such arrangements will by itself remove neither the *causes* of disputes, nor those underlying factors on the side of firms and unions which have made the industry—and particularly one section of it—notably liable to strikes.

The strike-wave that stimulated the joint meetings came to a climax in the late 1960 recession in the car trade. Analysis indicates that this represented a peak in the car firms' strike-incidence. If, as output recovers, the motor industry's "loss of working days" from disputes falls in 1961, it will thus still be rather early to assess the joint meetings' effect.

compared with the immediate post-war years and on our estimates, had multiplied threefold by 1957/58 and sixfold by 1959/60. While, measured in the conventional terms of the reported "loss of working days" from industrial disputes as a ratio to the number of their employees, the liability of the car manufacturers to strikes now rivals or exceeds that of such traditionally strike-prone industries as coal mining, shipbuilding and the docks.

This was perhaps not altogether to be expected, since another index of industrial discontent, labour turnover, has been distinctly lower for the motor vehicle industries than for manufacturing generally (16 per cent. of terminated engagements against 25 per cent.,[1] for men employees and in 1959) ; and the car workers in particular have no such inter-war history of protracted under-employment and wage-depression as that which apparently confirmed the strike-habit in older industries. To the contrary, in fact. The motor industry has been one of the more dramatic growth sectors of this century's economy. And although pay-packets in the car factories have been of recent years a matter of envious popular reference, wages in the car firms have always been high.

That the increase in the motor industry's strike-incidence was also not altogether inevitable might be thought suggested by the remarkable difference in its distribution between firms. The apparently-absolute freedom of Vauxhall's from the stoppages that have troubled other members of the so-called "Big Five" is now well-known : but similar contrasts are to be found among the smaller concerns. Some of these appear to have experienced a greater liability to strikes than major manufacturers ; while at the other extreme we have not traced a single stoppage among the car-workers of Rolls-Royce.

The Car Firms and The 'Motor Industry.'

Since nearly all the figures in this paper will represent the writers' own estimates, it would be as well to begin with a few definitions. There is no official series for industrial disputes in the motor industry over the post-war period. Ministry of

[1]Estimated from *Ministry of Labour Gazette*, quarterly reports of "Labour turnover rates in manufacturing industries."

Labour strike-statistics up to 1948 included the motor trades in the general engineering industry : from 1949 to 1958 they were grouped in a vehicle industries category which included aircraft manufacturing. For the past two years, 1959/60, the published figures relate to a "motor vehicle and cycle" industry group which embraces not merely the car assembly factories but the specialist makers of heavy road vehicles, of two-wheelers, three-wheelers and other perambulatory devices, plus a great number (but not all) of the component and accessory manufacturers, as well as various repairers and other establishments which are included in the category on technical grounds. But to trace the growth of recent industrial unrest in the motor industry one must go back at least to its reconversion from war production. While many of the establishments included in the current grouping (which follows the Standard Industrial Classification) seemed irrelevant to this purpose : their techniques and structure differ from those of the car firms which constitute the group's nucleus ; so may their economic experience—for instance, during the 1960 car recession the heavy vehicle makers experienced a continuing boom.

But in any case, the problem of labour unrest in the vehicle industry seems highly concentrated in the car firms proper. We are therefore concerned only with establishments that actually assemble cars and light vans, with their associated engine and transmission plants : except that some of these make their own bodies, some have them made by subsidiaries, while others buy them from outside. The Pressed Steel Company, for instance, supplies several car-manufacturers with bodies fully-assembled, painted, and trimmed, and awaiting only their working parts : we have therefore included the concerns which make car bodies in our analysis. We thus have a compact group of firms and establishments, relatively homogeneous in techniques and economic context, in which the Big Five are responsible for some 80 to 85 per cent. of the employment and over 95 per cent. of the final output.

The effect of selecting this group from the broader Ministry of Labour category—and the concentration in it of the vehicle trades' strike-incidence—can be seen by comparison. During

1959 and 1960 employment in the official "motor vehicle" industry averaged about 430,000 and the total reported "loss of working days" from industrial disputes was 830,000 : the total employment of the car plants as we define them was probably less than half the former figure, but we estimate their "loss of working days" for these two years to be some 740,000. Outside the car-assembly firms proper, in other words, the "motor industry's" dispute-incidence is in no wise remarkable.

More Definitions : and Sources.

But what is a dispute, for this purpose? In estimating the strike-liability of our "car industry" in general, as well as the development and causation of strikes there at large, we have been able to use not merely the summary accounts of "principal stoppages" which the Ministry's *Gazette* publishes each month, but the similar summary records of lesser disputes which are otherwise only included in the Ministry's statistical tables. Apart from one or two Courts of Enquiry reports on major disputes, our other continuing source of information has been the press—on which we are particularly reliant for detailed accounts of individual strikes and thus (since official assistance does not permit this identification) for our assessment of the position of individual firms and plants within the general picture ; since newsprint rationing ended, at least, public interest in strikes generally and cars in particular has ensured a press coverage not very dissimilar from that of official reportage.

But there is no obligation on employers or trade unions to report disputes ; and official reportage, particularly, excludes strikes below a certain minimum size in any case. Thus very many minor stoppages may be included in neither our statistics nor the Ministry's. For instance, the firm of Briggs' Bodies complained in April, 1957, to a Court of Enquiry into a dispute that in the previous eighteen months it had experienced 234 stoppages : over the same period, so far as we could trace, only a dozen disputes at the firm were reported in the press or big enough to be included in the official statistics. Similarly, Ford's protested to the trade unions early in December, 1960, that its Dagenham plant had had 57 strikes in the year : of

these, only ten appear to have been publicly recorded. While apart from very small strikes, such alternative forms of workers' protest or pressure as "go-slows" or overtime bans are not officially recorded at all, except in so far as they develop into an open stoppage of work.

These things make it undesirable to introduce into estimates of the present kind material from sources other than those referred to : since private information could only relate to individual firms, sections of workers, or periods, its use would bias comparisons. However, if public reportage of industrial disputes represents only a large sample of actual "dispute-incidents," it *does* cover by far the greater part of "working days lost" by plants where disputes occur. This is because of the concentration of that loss in the larger stoppages, which are always widely reported. Thus, even excluding the effect of two or three national stoppages which have involved certain car manufacturers, a bare dozen biggish strikes at individual firms are responsible for over half the car plants' estimated total post-war "loss of working days."

But this last term itself raises a question of meaning. The word "loss" is, of course, loaded (though it represents the official terminology) : output held up through disputes may be regained through overtime or increased effort afterwards—there is even a respectable argument that the monotony of work in some trades is such that the occasional drama of a strike helps to maintain labour morale and productivity. In the motor industry particularly, however, larger disputes have often occurred during periods of recession, when what has been "lost" would not otherwise be wanted anyway : had not the 11,000 workers who were involved in the Rootes' stoppages of October/November, 1960, for instance, in fact struck or been laid off in consequence, the 80,000 days thus "lost" would only have meant still shorter short-time all round.

Apart from the matter of terminology, however, there are three possible measurements of this "loss." In any stoppage (which for practical purposes, these days, means a strike—lockouts being almost unheard-of) one can take the loss by the workers who actually strike in the first place. Alternately, one can take the loss by the establishment or plants actually

struck, including that on account of employees laid-off because no work is now available to them. Or finally, one can take the loss by *all* the workers stopped in consequence of the dispute at all the establishments in some way, however indirectly, affected by it. Because of the elaborate division of labour and integration of production in the motor industry (so that very small disputes may cause very large interruptions of the work flow) the difference between these three measurements may be particularly great in the car firms. One B.M.C. strike of less than 200 crane drivers, for instance, stopped 6,000 other workers in the same plant, and was then reported to have led to 8,000 operatives in other B.M.C. factories alone being laid off.

In theory, the first of these three measurements is most useful if an index of "industrial militancy" is required : the third should be the best index of strikes' economic effect. In practice, the first measurement may be misleading because workers only "indirectly involved" (whom it excludes) might have been laid-off to put pressure on the original strikers, or to forestall a refusal to handle "black" work or materials; if they had not been laid-off, they might have "come out in sympathy"; they might even—as in the case of certain dayshift workers in Pressed Steel and B.M.C. during the protracted "four-night week" dispute of summer 1960—have otherwise struck against having their work interrupted by strikes ! In practice, again, the third measurement (and aside from the objections to its validity as a "wastage" index) is rarely available.[1] We have therefore compromised by using the second measurement (which corresponds approximately to the Ministry of Labour's "loss of working days by workers directly and indirectly involved") which in any case usually seems the one most accurately reported.

The Overall Strike-Incidence in Car Firms.

With all these preliminaries and qualifications, we can summarise the effect of our analysis so far as the total incidence

[1]However, Mr. F. Beswick, M.P., has stated (presumably from an official source) that for half-a-million days reported lost by the whole "motor industry" in 1960 at establishments involved in disputes, a further 300,000 were lost in consequence at other establishments. (*The Autocar*, 21st April, 1961).

of strikes in the car firms and its increase during the post-war period are concerned :

TABLE I

POST-WAR INDUSTRIAL DISPUTES IN THE CAR FIRMS

Annual Average of 3-year period	Number of Separate Strikes Traced	Workers "Directly and Indirectly Involved"	"Working Days Lost"
1946–48	12	15,000	75,000
1949–51	13	17,000	45,000
1952–54	10	20,000	185,000
1955–57	25	40,000	135,000
1958–60	62	105,000	285,000
Total, 1946 to 1960 inclusive	369	—	2,150,000

Currently, and so far as we can reckon it, the total employment of the car firms proper is not much more than 200,000, while over the whole post-war period it has probably averaged between 150,000 and that figure. Allowing for the additional effect of several general engineering stoppages in which car-workers have been involved (*see below*) the incidence of strikes in the car firms, when measured in terms of the ratio of working days lost to the number of employees, has thus risen from about four times the national average in the immediate post-war years to some seven or eight times the national average in the three years up to 1960. And since in the car firms the ratio of "staff" employees—who have struck, but only on two or three quite recent occasions—is higher than in other manufacturing industries, the disproportion in the car-makers' strike incidence would probably be rather greater if it were reckoned only for their 150,000 or so manual workers.

At the same time, it is as well to keep matters in perspective. Estimating on the same basis as Table I one can put the total involvement of car factory employees in strikes since the war

at about 600,000—counting any worker who struck twice in that time as two, and so on. Which implies that the average assembly operative has been involved in only four strikes (or five, with the general engineering disputes referred to) since 1945, each of which lasted less than four days. While even in the last few years, the same average operative has lost—again in terms of the definitions—two days annually from disputes.

The estimates of Table I do not include some 500,000 days probably lost in those car factories involved in the three national strikes against "federated" engineering firms since the War—the One-Day Strike of 1953, the week's stoppage of 1957, and the apprentices' strike of 1960. A word on collective organisation among the car-making employers and workers might be in place here.

This could perhaps be best described as on the whole strong, but patchy and disjointed. On neither the employers' nor the union side is there a widely-recognised organisation specifically and independently concerned with the car or vehicle industry as such. The three British-owned members of the Big Five, and several of the smaller firms, belong to the Engineering and Allied Employers' National Federation. So far as wages go, this means little more than that their workers benefit by the national advances conceded for the general engineering industry and by its minimum "guaranteed week." Most of their production workers are paid piece-rate or other bonuses far above the basic rates agreed by the Federation, whose right to control such payments has been resisted. Even in "federated" firms, basic hours are sometimes shorter than those nationally-agreed. Such firms are, however, bound by Federation "dispute procedure" and certain other rules (like those reserving "managerial functions"). Rootes and one or two other firms also recognise certain agreements of the Joint Wages Board for the Vehicle Industry (which deals mainly with conditions in non-car establishments). Ford and Vauxhall belong to no employers' association or external negotiating body.

On the workers' side, all the unions, except one or two "breakaways" of dubious membership (like the Toolmakers') affiliate to the Confederation of Shipbuilding and Engineering

Unions. Conversely, most of the C.S.E.U's two-score affiliates have members in the car firms, though only one of them is primarily concerned with the vehicle trades. All firms have a strong union membership, but particular departments or sections may be badly-organised : there have even been strikes by operatives who were initially non-unionists—for instance, certain earlier disputes among internal transport and delivery drivers. There is no systematic—or even widespread—division of interests between the unions, but the four largely concerned with assembly workers and their like are the two "general unions" (T. & G.W.U. and N.U.G.M.W.), the A.E.U., and the National Union of Vehicle Builders. The last two also organise skilled operatives, but several craft or "near-craft" unions—the Sheetmetal Workers, the E.T.U., the Draughtsmen, and so on—have important memberships in car plants. The shop stewards of the various unions have joint organisations among themselves which are "officially" approved or tolerated in roughly inverse proportion to their coverage.[1] In general this complex structure of collective organisation clearly derives from the car firms' origins as merely specialist engineering concerns.[2]

The Causes of Strikes.

Some people think the explicit causes of strikes to be unimportant, regarding them merely as accidents that spark off a situation created by deeper factors. But the differences in strikers' demands (or protests) between industries, firms, and periods are too systematic for the immediate issues in strikes to be considered anything else but highly significant. This is not to say that the background condition to a high strike-propensity can be ignored. The association between production changes and strike-movements is too intricate to

[1]Thus, the T.U.C. General Council's report to its 1960 Congress on strikes justified joint shop-stewards' committees at the plant, but condemned general shop stewards' movements. Another interest in this level of car-industry unionism is perhaps indicated by the 1961 Communist Party Congress's election of a notable B.M.C. convenor, Mr. R. Etheridge, to its Executive.

[2]There is a more detailed description of collective organisation in the vehicle trades in a previous study by one of the present writers ("The Crossley Strike," *Manchester School*, September, 1950).

examine here[1]: but in car manufacturing as a whole, for instance, output has multiplied six-fold since 1946, so that some might think the parallel rise in the industry's strike-incidence not unconnected with the strains of rapidly increasing labour productivity; and we hope later to indicate something of the contexts that have induced a particularly high dispute-liability in certain firms. But for the present, it is enough to say that the immediate issues in strikes are usually of a kind that might very well represent to the people concerned grievances sufficiently real to explain their stopping work about them.

At any rate, the causation of strikes in the car firms seemed sufficiently interesting to justify an attempt to analyse it in some detail. The Ministry of Labour, of course, classifies the disputes in its records according to their causes, and following as best we can the Ministry's method of classification (only a little simplified, since some types of stoppage, like pure "sympathy" strikes, have not been sufficiently numerous in the motor industry to justify separating them) we get the distribution shown by Table II.

The relation between the Table's two sets of figures incidentally tells one something about the size or obstinacy of different kinds of strike. Thus, strikes for wage-increases tend to be bigger than the norm (which, in the car firms, can be taken as a stoppage of some 1,500 workers lasting about four days): this is partly because we have here included an estimate of the effect of the three general engineering strikes in which "federated" motor concerns were involved. But strikes about dismissals also tend to be big—and so, outstandingly, do those about trade union questions (like the "union shop," "jurisdiction," or the status of shop-stewards).

In fact, of the dozen stoppages already referred to as responsible for over half the car firms' post-war "loss of working days" from internal disputes, no less than eight concerned dismissals or trade union issues. And these eight included not only such recent affairs as the Rootes "redundancy" stoppage

[1]This we discuss in a separate article, "Strikes, 'Redundancy' and the Demand Cycle in the Motor Industry," *Bulletin of the Oxford Institute of Statistics*, May, 1961.

TABLE II.—CAUSATION OF STRIKES IN THE CAR FIRMS, BY OFFICIAL CATEGORIES, 1946–1960 INCLUSIVE

Percentage attributable to :	Wage Increase Demands	Other Wage Issues	Hours of Work	Dismissals	Other Employ-ment Issues	Other Working Arrange-ments	Trade Union Issues and Sympathy
Of all *Strikes* :	30	17	2	16½	8	15½	11
Of all "*Working Days Lost*" :	38½	10	2	19	3	10½	17

TABLE III.—CAUSATION OF STRIKES IN THE CAR FIRMS (AS RE-CLASSIFIED) 1946–60 INCLUSIVE

Percentage attributable to :	Straight Wage-Demands	Wage-Structure and Work Loads	Redun-dancy, Short-time, etc.	Individual Dismissals, etc.	Manage-ment Questions	Working Conditions and Hours	Trade Union Relation-ships
Of all *Strikes* :	12	38	14	9	6	5	17
Of all "*Working Days Lost*" :	16	24	24	5	3	3	25

of 1960 or the 1959 "closed shop" strike at Ford, but much earlier conflicts—beginning in 1946 with one at Daimler's, and running through a sequence of strikes at the Austin plant and in Dagenham over 1951/53 that alone "cost" more than half-a-million working days between them. On the other hand, stoppages on "other wage issues" (which include the system of payment, allowances for work under non-standard conditions, earnings guarantees and the like) though apparently quite frequent, are generally smallish. So are strikes about "other employment issues" (which are mostly demarcation or dilution disputes) and those about "other working arrangements"—which is virtually a miscellaneous category.

The main purpose of the table, however, is to permit comparison with what is known of strike-causation in industry generally. So far as the frequency of different kinds of dispute is concerned, the proportions of strikes attributable to the various causes distinguished in Table II do not seem very different from the recent distribution in manufacturing industry at large—except that dismissal disputes seem, perhaps surprisingly, rather less frequent in car firms than in other industries, while "trade union issues" seem a notably more common cause of strikes in the motor factories.

When we compare the causation of "working days lost," however, not only trade union issues but dismissals emerge as a much more substantial source of lost time, relative to other causes of strikes, than in other industries—which emphasises the size of car plant stoppages on these questions. And these two causes of dispute are often in fact associated : several of the biggest post-war strikes, for example, have arisen from the dismissal of individual shop stewards or their inclusion among groups of workers discharged.

The importance of disputes on trade union questions in the car firms does not seem hard to explain. The structure of the engineering unions involves implicit rivalries of two kinds : between unions, for membership and bargaining strength : and within unions, between the national officials and the shop-stewards as alternative foci of leadership. In longer-established fields of engineering union interest, these rivalries have become more contained. But the rapidly-expanding motor industry's

influx of new workers represented potential recruits to competitive unions and leaderships. While its management has numbered several forcible "personalities" who, if not quite technocratic in outlook, have been sometimes impatient of the human paraphernalia of production and inclined to regard the workplace union organisation in particular as a challenge to their authority and competence. On the other hand, two apparent factors in Vauxhall's freedom from disputes are that it has itself gone to some trouble to provide (in its elected and active Management Advisory Committee) an effective form of workplace consultation, while one major union has been dominant among its workers.

But the analysis of Table II involves a certain confusion. The official strike-category which corresponds to dismissals as a cause of dispute includes not merely the discharge of individuals on disciplinary or similar grounds, but the block dismissal of operatives for whom work is no longer available. On the other hand, a strike around a demand for "work-spreading" *instead* of a threatened redundancy might come to be officially classified as one about either working hours or "other working arrangements" according to the method proposed—or even, if that method included an "earnings floor"—as being about "other wage questions."

A similar problem, however, arises in regard to wage-disputes themselves. In most industries, demands for wage-increases appear to represent the biggest single source of "days lost" from disputes : but it is not easy to understand why this factor should be quite so prominent in the car firms' strike-record. For a large part at least of the post-war period wages for normal hours in the major car plants have probably been about half as much again as those ruling in engineering.[1] Just before the 1960 recession's onset, men's average weekly earnings in the "motor vehicles" group of manufacturing industries together were £19—more than £4 above the general

[1]From an estimate by one of the present writers (in "Wages : Industry Rates, Workplace Rates, and the 'Wage-Drift,'" *Manchester School*, May, 1956). See also the data of Knowles & Hill in "The Structure of Engineering Earnings," *Bulletin, Oxford Institute of Statistics*, September/October, 1956.

manufacturing average for the same hours worked[1]; while in the assembly plants of the "Big Five" car-makers they probably averaged about £2 to £3 above the figure for the motor trades as a whole, despite a lower proportion of skilled workers in the car factories. Yet even excluding the national engineering strikes' effect, stoppages in support of wage-increase demands would—according to Table II—apparently account for a quarter of the car firms' post-war strike-liability.

On examination, however, it becomes clear that very many of the strikes officially classified as due to demands for wage-increases in fact spring from differences in rates or earnings between related jobs in the same plant or firm. B.M.C. workers at Llanelly have struck because their wages were lower than those in the firm's main plants, and Pressed Steel operatives at Swindon because they thought their earnings lower than in Oxford; and within individual plants claims arising from differences in piece-workers' earnings, or between those of production operatives and auxiliary workers, are a very common cause of disputes. Such disputes, in fact, seem more concerned with the *structure* of wages than their absolute level.

The Major Sources of Dispute.

Inevitably, classifying strike-causes is an exercise in arbitrariness. At B.M.C's Tractors and Transmissions plant, 120 toolroom workers struck because fourteen fitters had been suspended for refusing to work an extra shift when other operatives were on short-time. Is this a strike about hours, discipline, or employment—or a "solidarity" stoppage? But arbitrariness can be more, or less, relevant. It is not a criticism of the Ministry of Labour's method of classifying strikes to say that it has been historically established, in the light of the experience of older industries and antique strike-movements. But we do not think it the most revealing way of analysing the motor firms' strike-causation.

We have therefore re-examined the actual causes of all the strikes of which we have been able to discover some record, so as on the one hand to group together all those originating

[1]cf. *Ministry of Labour Gazette*, "Earnings and Hours in April, 1960."

primarily from a state of work-shortage or labour surplus (thus distinguishing them from purely individual dismissals or questions of shop-stewards' status), and on the other hand to separate strikes for "straight" wage-demands from those arising from the structure, system or composition of wages. It is often said, however, that some firms (particularly Ford and Vauxhall) avoid the disputes associated with piece-rate fixing by paying only a time-based wage : but among time-workers an analogous source of trouble exists in the workload to be accepted in return for the fixed hourly wage (expressed, for instance, in the not infrequent stoppages about the size of time-paid working gangs). So again we group such disputes together with those about piece-rates and wage-differentials as concerning "wage-structure" or related questions.

When these changes are made, all that is left in the official category of disputes about "other working arrangements" is two distinct types of stoppage ; strikes about the rights of management—for instance, to transfer workers or determine their supervision and working methods ; and strikes against physical working conditions—heat and ventilation in the paint-spraying shop is a hardy annual. Strikes about hours similarly boil down to some recent ones about shift-work (like the "four-night week" stoppages of 1960, or those against three-shift working in Ford's) which are too few to justify separating from others about conditions of work. While the official group "other employment questions" has little left in it but a recurrent disagreement on work-allocation between members of the N.U.V.B. and the Sheetmetal workers, which can be grouped under "trade union issues" with other types of inter-union dispute. The result is shown by Table III. The overall picture is immediately rather different to that presented by Table II. Here we have not included the effect of the general engineering strikes. But even if we had (in which case "Straight Wage Demands" would account for 30 per cent. of "Working Days Lost") wage-structure issues, redundancy, and trade union questions would each have accounted for one-fifth all time recorded as lost from disputes.

Again this says nothing about the *general* validity of the official method of classifying the causes of strikes. But in the

particular circumstances of the motor industry, its effect is not merely to understate the impact of what we have already distinguished as one major internal source of dispute—trade union relations—but to very largely conceal that of two others : "redundancy," and the industry's wage system and structure (or rather, perhaps, in view of the negotiating arrangements already described, its lack of those attributes so far as wages are concerned).

It is, in fact, these two issues that have been particularly responsible for the marked increase in the car firms' strike-liability during the past few years. The contribution of the three major causes of strikes to this movement can be shown by contrast with the immediate post-war decade, as in Table IV.

TABLE IV

DEVELOPMENT OF STRIKES IN THE CAR FIRMS, 1946–60 INCLUSIVE.[1]

	Number of Strikes attributable to :	Lost Working Days (000's) attributable to :		
	Wage-Structure and Work-Loads	Redundancy, Short-time, etc.	Trade Union Relationships	All other Causes
Annual Average of 1946–55 :	4 (17)	2 (15)	3 (34)	3 (30)
Annual Average of 1956–60 :	19 (69)	7 (76)	7 (38)	15 (60)

[1]Again, the "days lost" estimate does not include the effect of general engineering stoppages : with these in, the annual average loss from "all other causes" would probably be about 40,000 days up to 1955, and about 150,000 from 1956.

Each of the three major strike-justifications distinguished there accounts for over a half-million "days lost" since the War. But the incidence of "trade union" disputes has not much increased in recent years, compared with that from other causes. And, from internal sources at least, it is wage-structure and redundancy issues that are largely responsible for the car firms' increased "loss of working days."

Changes in Company Structure and Union Policy.

But this brings us to the apparent turning-point in the car workers' dispute-history—the 1951/52 recession.[1] Why, for instance, should redundancy have become so outstanding a cause of strikes in just the industry where growth has been swiftest? The post-war period has, of course, been one of continued technological change. The psychological impact of this has been perhaps inflated by the car factories' particular identification with "automation"; but at least this extended mechanisation from skilled operations to many unskilled jobs like parts handling—so that in recent years growth of the manual labour force has virtually ceased. The period has certainly been one of change and concentration in the car firms' economic structure. Though this, too, continued an inter-war trend, it brought the industry near to the point at which output might be foreseeably concentrated in the hands of two or three major producers, and even the survival of certain members of the "Big Five" themselves has seemed in doubt. Relatively old-established firms like Jowett and Singer have disappeared or been absorbed in others; Morris and Austin have merged in B.M.C.; Standard has been "taken over"; once-independent body-makers like Briggs have been amalgamated into the final assembly firms.

Such structural changes have been accelerated by the short but severe slumps that have interrupted the otherwise rapid post-war growth: the recession of late 1960, in which output fell by half in a few months, was the third in a decade, so that the industry's progression had already begun to assume an abrupt cyclical pattern. The extraordinary post-war expansion, with output rising 25 to 30 per cent. annually in non-recession years, was at least half-based on exports, in which the British industry had a temporary special advantage. Since the recession of 1956, overseas competition has become intense, and the annual output growth has flattened to between 15 and 20 per cent. Nevertheless, the announced investment plans of the major manufacturers would add another 50 per cent. to the industry's capacity, so that some think these plans only

[1]Cf. the last column of Table I.

reconciliable with the economic trend through a mortality among the car firms themselves.

Against this background it is perhaps understandable that insecurity should emerge as a major theme in car workers' disputes. However, it is not so much the increased *frequency* of "redundancy strikes" that accounts for their recent importance, as Table IV indicates. The remarkable thing about them is rather (as comparison of the Table's two halves will show) that whereas other kinds of strike have tended to get smaller as they became more frequent, stoppages on redundancy and related issues have become bigger—so that they currently represent the most substantial sort of car-industry dispute. Now the critical determinant of a strike's size and duration is usually the extent of support it receives from the unions. Overwhelmingly, car firm disputes appear to have begun "unofficially" : but since 1952—witness the T. & G.W.U. strike for "redundancy compensation" against B.M.C. in 1956—the unions generally have been much more willing to support (or to approve in retrospect) stoppages on this issue at least.

However, "wage structure and workload" disputes not merely rival redundancy issues as a source of lost time over the whole post-war period : they constitute by far the biggest single cause of individual strikes, well over a third of which have been attributable to them (*see Table III*). And wage-structure issues are especially responsible for the remarkable rise in the frequency of disputes, from an average of 13 annually before 1956 to one of 48 annually from that year on : over two-fifths of the increase arises from this source. While, unlike the redundancy disputes, which become more important in the slumps in the car industry's activity, "wage-structure/workload" strikes become more frequent in booms, and are thus considerably more disruptive of production.

Moreover, even our own classification may somewhat understate the impact of "wage-structure" as a factor in disputes, since we have included in strikes about "straight wage-demands" not only those explicitly and primarily based on the prosperity or productivity of the car trade or particular firms, on the cost-of-living, and the like, but also wage-disputes about which we had no detailed information. While even some

of those we have deliberately included in this group might well be argued to be in fact about "wage-relativities." For instance, the first substantial strike after the War's end was for a wage-increase at the "non-federated" Dagenham firms in August, 1945 ; and had we included 1945 in our figures (we didn't, although the War's conclusion was apparently greeted by a minor strike-wave, because many car plants had still not resumed their normal work) we would have treated it as due to a "straight wage-demand." But it arose from an arbitration award to workers in "federated" concerns, and might equally be regarded as one about wage-structure.

If the rising incidence of redundancy strikes has been largely traceable to the changing attitude of the unions, the extraordinary upsurge of wage-structure disputes seems clearly connected with changes on the side of the car firms themselves. Up to 1953 the annual frequency of such strikes was remarkably regular, at just about four each year. In 1953, however, they virtually ceased. But over 1954/55 they re-appeared, their frequency accelerating—with a brief interruption by the 1956 recession—until over 1959/60 they were responsible for nearly half of all stoppages and some 40 per cent. of "days lost." And this, despite the fact that they include hardly any really big disputes with substantial "official" backing. So that, in sum, the wage-structure strike-wave appears to represent an almost wholly spontaneous movement on the operatives' part.

But it was during the 1951/52 recession that the major post-war changes in the car firms' own structure were initiated. The merger of Austin and Morris, with the latter's subsidiaries, into B.M.C. was announced in 1952 ; Ford bought Briggs' Bodies in 1953 ; and the Rootes Group acquired Singer in 1955. The whole pattern of strikes in the car plants is in any case heavily influenced by events in particular firms. For instance, the fact that even excluding the general engineering stoppages strikes for "straight wage-demands" remain larger than the average (*see Table III*), is mainly due to their concentration in certain big "non-federated" establishments. But the coincidence of the "wage-structure" strikes' acceleration with these other

structural events warrants an attempt to examine the strike-liability of particular concerns more closely.

THE STRIKE PATTERN BY FIRMS.

The change in the bigger firms' strike-proneness in general can be seen by comparison between the experience of the "Big Five" together and that of the independent concerns. In fact, of course, the term "Big Five" is now rather a misnomer. Two of the block's members currently finish nearly 70 per cent. of British cars, while Standard's relative position has declined. Its share in car assembly work is now probably rivalled by Jaguar—which has recently absorbed Daimler—and exceeded by Pressed Steel, while among the handful of other "Independents" remaining, Rover is outstanding. However, the demarcation is still accepted and would be valid for most of the post-war period : while even with only indirect help from official sources, it is possible to separate the two groups' dispute-records.

It has been suggested that the size of firms is itself a factor in industrial unrest, big ones being more liable to strikes than little. Our study offers no evidence to support this. In the first five post-war years, the Independents actually had rather more strikes than the "Big Five," and were responsible for more than a quarter of "days lost"—a figure which, even then, was disproportionate to their share of the car firms' total employment. While over the whole post-war period, exactly a third of all the strikes traced have occurred in independent plants. This includes the experience of certain firms—like Jowett—which have ceased car-production during that time. But the more substantial Independents—Jaguar, Rover, and Pressed Steel—each seem to have had about as many stoppages as Standard, which has been thought one of the more strike-prone larger firms.

There is nothing outstandingly different about the pattern of dispute-causation in the independent firms as a group which would explain their disproportionate strike-liability over much of the post-war period—and particularly, their greater frequency of stoppages. So far as the latter is concerned, the percentages of their strikes attributable to different causes

approximate to the distribution shown in Table III for *all* car firms together. Their share of "wage-structure and workload" disputes is rather higher : but this is probably due to the fact that practically all the independent firms use "payment by results" wage-systems, unlike certain members of the "Big Five." Again, once disputes originating in redundancy are separated from strikes against individual dismissals, the latter seem notably more frequent in the smaller firms—which are responsible for nearly half the stoppages we have traced on this issue. This may be due to the fact that firms like Rover and Jaguar have a much higher proportion of apprenticed craftsmen than the big low-price producers ; and these workers have a traditional group-solidarity which is not so developed in the general run of assembly operatives.

When one considers the distribution of "lost working days" between the various causes of dispute, however, the position of the Independents is rather different : even without the general engineering strikes (which, proportionately, speaking, probably affected them rather more than the "Big Five") about half of all their post-war "days lost" is due to wage-disputes, against less than two-fifths in the major concerns. This also implies that wage-strikes tend to be relatively more stubborn and widespread in the smaller firms. It rather looks as if a major source of the Independents' strike-liability has been their strongly-organised craftsmen's resentment of high earnings by the "popular quality" operatives of the big plants. Jaguar, for instance, is in the same district as Standard, where average wages have been reported the highest in the industry.

But the strike-liability of the Independents has not been rising so fast as that of the "Big Five" : in the last five years they have experienced 30 per cent. of the disputes traced but only one-sixth of the "days lost"—the last figure being just about proportionate to their current share in car workers' employment. And that they have even maintained this position—as a group—seems substantially due to Pressed Steel, the fortunes of which largely follow those of the "Big Five's" British members, which it supplies. Until 1955 this firm seems to have been relatively strike-free : but since then it has experienced quite a few disputes, including some biggish ones.

It may in fact have become involved in B.M.C's strike-movements (its main plant, of course, neighbours the Morris works at Cowley).

The recent increase in the car-firm's strike-incidence is thus largely traceable to members of the "Big Five." In analysing the position of individual firms, one can of course not appeal to official strike-records : but even were not such very extensive press reference available, a little knowledge of the motor industry suffices to identify most of the plants *Gazetted* as suffering from "Principal Stoppages," and the latter seem responsible for some 85 per cent. of all days reported as "lost" from car industry disputes. At any rate, the post-war strike-experience of the "Big" firms can, we think, be not unfairly apportioned between the four which have some (Vauxhall having none) as follows :—

TABLE V

STRIKE-EXPERIENCE OF MAJOR CAR FIRMS, 1946–60 INCLUSIVE

Percentage (for all car firms as defined)	British Motor Corporation	Ford	Rootes Group	Standard Triumph International
Of all *strikes* :	37	14	10	7
Of all "*working days lost*" :	44	23	8	9
Of current *employment* (approx.) :	35	25	8	5

For each combine we have included the experience of all concerns which have merged in it since the War, as well as an estimate of its loss from general engineering strikes. But one shouldn't infer *too* much from comparison of the lower two lines, since Ford's employment has been expanding relatively and Standard's declining. Similarly Ford's relatively low percentage of *strikes* partly means that its stoppages have been individually much bigger than the norm, while those in Rootes plants have been smaller.

THE BIG FIVE : INCLUDING THE BIGGER TWO.

The cluster's junior members can be covered briefly. Despite Standard's reputation, it appears to have had no disputes worthy of report before 1954. From plants now in the Rootes Group, however—which until very recently had a peaceable repute—we have found a pretty steady two or three yearly throughout the period covered. But in 1954 Standard had a biggish strike against the dismissal of several shop-stewards, followed in 1956 by a very big one arising from redundancy. These two stoppages—which seem to have followed changes in managerial policy [1]—account for most of S.T.I's "loss" of working days, and since then its strike-incidence has been very modest. Rootes' strikes, on the other hand—though they reflected the general increase in frequency during 1959/60—remained smallish until the last year : but in October the operatives at nearly all its plants walked out in support of an anti-redundancy strike at its Thrupp & Maberly subsidiary, a second sizeable stoppage following in November. So that Rootes' "strike-loss," too, is highly concentrated. Both these firms have had a particularly large share of redundancy disputes, apart from the two big ones referred to : perhaps because they have each produced mostly for a middling price-range which may be most sensitive to car trade fluctuations. Rootes, however, is notable for a largish proportion of "wage-structure" stoppages—possibly a combined effect of its traditional wage-system and the dispersal of its plants—but unusually good trade union relationships (to judge by its freedom from overt troubles on this score). S.T.I., on the other hand again, has had several disputes on questions of management, which in *most* car firms are rather rare.

Two-thirds of the car firms' total "strike-loss", however, has fallen on what might be called "The Bigger Two." And between Ford and B.M.C. the contrast is more marked. Ford has a very simple national wage-structure based on time-rates—though its low proportion of "wage-structure" disputes is partly offset by its liability (being "non-federated") to direct

[1]Standard introduced its "large gang" working system (see Melman, "Decision-Making and Productivity," *Blackwell*, 1958, for a description) in 1951. It was abandoned after the so-called "automation" strike of 1956.

wage-demands. But it has apparently had a relatively outstanding proportion of disagreements about "management questions," which together with unusually frequent disputes over individual dismissals and the arrangement of working hours, etc., make up two-fifths of the Ford strikes we have traced. While this is also associated with an abnormal frequency of troubles arising from trade union relationships. Which adds up to the Ford picture (or rather, Anglicised mirror-image) of a thrustful, competition-minded management which has only gradually accepted the restraints of collective negotiation and consultation, and still rejects the full complications of the British engineering trade unions' system.

But the two main Dagenham plants—Ford proper and Briggs'—were both prominent strike-performers before the 1953 merger, their shop-stewards co-operating "unofficially." And despite the troubles—culminating in the famous "bell-ringing" strike of 1957—involved in Briggs' assimilation to the Ford system, none of these matched the great Briggs-Ford stoppage of 1952. If one compares the experience of the combined group up to and after the merger, its strike-frequency has risen less than the average since 1953 and its annual "loss of working days" actually seems somewhat down. So that—although by the standards of other trades Ford is still exceptionally strike-prone—the firm has not contributed much to the more recent uprush of car-industry strikes.

To this uprush, however, there seems little doubt of the B.M.C. plants' contribution. The point can be put quite tersely, again, by a comparison of dispute-liabilities before and after the B.M.C. merger :

TABLE VI

POST-WAR DISPUTES IN B.M.C. AND OTHER CAR FIRMS

Approximate Annual Average of :	Strikes Traced		"Days Lost" (000's)	
	B.M.C. Group	All other Firms	B.M.C. Group	All other Firms
1946–52 inclusive :	3½	9	20	70
1953–60 inclusive :	14	21	130	125

If anything, the comparison of "working days lost" probably rather understates B.M.C's relative share, since with the definition adopted here, the effect of stoppages will be less fully recorded in the more dispersed concerns. Shortly, both the strike-frequency and the annual "strike-loss" of the car firms as a whole have risen by over 180 per cent. since 1952. Nearly half the increase in the number of strikes and two-thirds that in the "loss of working days" is apparently attributable to B.M.C. itself.

Two things are quite outstanding in the B.M.C. strike-pattern. Of more than a hundred post-war disputes at plants now in the group where we have been able to trace the causation, nearly 45 per cent. arose from "wage-structure/workload" issues. While nearly a quarter concerned some aspect of trade union relationships. Considering the extent to which B.M.C. itself dominates the industry pattern shown by our Table III, these proportions are quite abnormal. "Wage-structure" disputes, particularly, appear responsible for half the group's increased strike-liability since the merger. B.M.C. retained, at the time of our enquiries, the system of individual plant management responsibility for labour questions, subject to intervention by the local Engineering Employers' Association in unsettled disputes : and a national trade union official could then complain of B.M.C. that " . . . with all this labour force, they do not employ one man of higher management standing whose full-time job it is to deal with labour relations in a co-ordinated way." [1] Clearly, the rapid technical integration that followed the merger was not matched in handling the labour problems it created. On the other side, however, the size and special dispersal of the group not merely exposed it to the full range of the industry's tangled inter- and intra-union relationships : the merger actually seems to have given the latter a new twist.

We have discussed the car industry's sectional strike-proneness from the angle of particular firms rather than that of particular districts, because local factors seem definitely subsidiary to intra-firm influences. The freedom of Vauxhall

[1] Mr. L. R. Kealey, National Officer of the T. & G. W. U., in the *Financial Times'* "Survey of the Motor Industry," October, 1959.

and Rolls-Royce from labour troubles has doubtless been furthered by their plants' geographical isolation from the main motor-industry centres as well as by their policies : Rootes' Dunstable plant seems to have been similarly strike-free. Between different firms' plants in the same district, strike-movements often seem to spread—as in the case of Morris and Pressed Steel in Oxford, or Ford and Briggs at Dagenham. But not invariably : the Rootes' Coventry plant was unaffected by the Standard dispute-wave of 1954/56, or by the previous militancy of Jaguar's workers, while Standard's factories did not obviously respond to the Rootes' movement of 1960. On the other hand, the latter crescendo seems to have been set off in London : British Light Steel Pressings, the Rootes' Acton subsidiary, seems to have previously experienced more strikes than other (and bigger) plants in the group—including a substantial one in January, 1960, and another in June. And each of the three firms with plants dispersed about the country has a "joint shop stewards' committee"—"unofficial," but not consistently condemned by the unions, being inevitable in view of their own structural hiatuses.[1]

Thus, Birmingham has been most prominent as a centre of car industry disputes not merely because of its grouping of several firms' plants, but especially because of the great Austin works' presence. Austin was the more concentrated partner to the B.M.C. merger. And whereas the dispersed Morris factories had a low strike-incidence before 1952, Austin had already had a post-war dispute record that included two complete stoppages—in 1948 and 1951. Part of the background to the great strike there that followed the merger in 1953—and which the Vehicle Builders' E.C. short-cut that union's rules to declare "official"—was the apprehensions the merger introduced at two levels of trade unionism. For instance, the N.U.V.B., which seems particularly strong in Birmingham and which had the one-time ambition to become an "industrial union" for the motor trades, was afraid that it would be over-matched, in dealing with the mammoth combine, by bigger unions—the gigantic T. & G. W. U. has a strong membership in the Morris

[1]See the discussion of the T.U.C. 1960 Report (*loc. cit.*) on committees of the B.M.C. type.

plants. And militant shop-stewards were afraid that the new combine would prefer to deal directly with the national union officers, thus reducing their own influence (as in fact happened to Briggs' stewards after the Ford merger). Austin men took the lead in the B.M.C. shop-stewards' "merger committee." And after 1953 the Austin strike-liability spread rapidly to the group's other plants—so that even though Austin militancy receded from prominence during one recent phase of that factory's management, the B.M.C. strike-pattern had by then become independent.

And a Conclusion.

Clearly, there are two levels of strike-causation in car manufacturing. To a significant extent, the sources of the industry's exceptional dispute-liability lie in particular firms. But the two or three firms which have not shared in it have enjoyed special circumstances as well as special policies. And at various times since the War different firms have been most prominent as trouble-centres—the "Independents" in the immediate post-war years ; Ford, Briggs and Standard in the early and middle fifties ; B.M.C. since (though outstandingly) and Rootes in the 1960 recession. The car firms' particular strike-proneness is at least partly due to factors affecting their trade as a whole.

So it is difficult to disentangle the possibilities of improvement at individual firms, either in labour relations generally or in several particular matters of dispute, from one common factor—the effects of car manufacturing's system of collective organisation. This seems apparent if one considers the three major causes of the industry's post-war "strike-loss"—wage-structure, redundancy, and trade union relationships.

Thus a settlement of the first two issues appears doubtfully furthered by most car firms' attachment to the engineering employers' federation. Although the cyclical fluctuations in car labour demand are not reproduced in any major section of engineering, the federation naturally fears that special "redundancy" agreements here would prejudice its other affiliates' position or the ultimate managerial right of "hire-and-fire." And despite the outstanding irrelevance of engineering's

basic rates as a guide to the car firms' actual wages—which are mostly determined by haphazard workplace bargaining—a decade's discussion within the federation has not yet produced proposals for a revision of the acknowledgedly-obsolete wage-structure of engineering itself.

Especially on redundancy, matters have not been helped by disagreements between a diversity of unions whose major interests lie mostly outside the car plants. But multi-unionism also makes conflict between *levels* of union leadership virtually inevitable, although one or two unions are attempting an improved internal co-ordination. When the workers of any plant or firm belong to several unions, with different rules (and sometimes policies), the necessary task of organising joint work-place bargaining often falls to the shop-stewards, who thus achieve an independent status.

In all this, a fundamental problem seems to be that the collective organisation of the old mechanical engineering trade has long outgrown itself : technical progress tends to make *all* industries engineering industries. In the U.S.A., Ford and General Motors (which in Britain are "non-federated") have found it possible to consult, together with other car firms, on labour contract revisions. Nor is it necessary to advocate "industrial unionism" to suggest that there is a case for an integrated organisation, whether within or apart from the present employers' and workers' associations, with sufficient autonomy to negotiate the distinctive labour problems of car manufacture itself.

JOHN BESCOBY,
Brussels and Sheffield.

H. A. TURNER,
University of Manchester.

Credit Restriction and the Supply of Exports

The United Kingdom is faced with the recurrent problem of reconciling a "satisfactory" rate of growth of income with an imbalance in international payments. A variety of instruments involving the use of fiscal and monetary policies, together with a battery of direct controls, are available for use in periods of crisis that are associated with a relatively high rate of expansion. The effects of such policies are usually discussed within the framework of general theories of the international trade mechanism involving the levels of income and relative prices, and the flow of short term capital.

However the more familiar picture of the adjustment mechanism has recently been overshadowed by the implicit assumption that a direct restriction of home demand will induce an increase in the supply of exports. Certain key industries such as the motor industry have found domestic demand restricted and their appeals for removal of restrictions turned down largely on the ground that restrictions have not been imposed sufficiently long to improve the balance of trade by increasing the export volume. Industry has countered this argument by asserting that a buoyant domestic market is essential to an expanding export market. A decline in home production it is claimed only has the effect of raising costs per unit of output, reducing employment and preventing the benefits of large scale production accruing to purchasers everywhere.[1]

The object of the present paper is to discuss these arguments in the light of two simple models of the firm. These models (like most economic models) are intended to be no more than suggestive and are not thought to be precise reflections of the behaviour of any individual firm. They serve to provide a useful framework for discussion and bring out several of the main points at issue.

[1]While it may be true that the motor industry is really being restricted in order to release labour, this is quite distinct from the argument that is examined on its own merits below.

I

I propose to consider the effect of the credit squeeze at the level of the individual firm. It is assumed that the firm produces a more or less homogeneous product part of which is exported and the remainder sold on the home market. Production is not allowed to exceed overall demand, although it may fall short of it. This means that the complicating factor of inventory changes can be put on one side. I assume also that subject to a limitation discussed in connection with the second model, the firm is interested in maximising profits in the short run.

The problem of determining the most profitable total output and its distribution between home and foreign markets depends on the market structure within which the firm operates, i.e. whether it sells under perfectly or imperfectly competitive conditions or adopts some version of full cost pricing.

A plausible model is that the firm is faced with negatively sloped demand curves in each market. It is assumed that the overall level of income in each country is held constant so that foreign demand depends only on price, and domestic demand depends not only on price but also on the degree of restriction that is imposed by the authorities.

This model reduces to the familiar model of price discrimination in economic theory. Maximising profits in the short run the firm will produce and distribute its output so that marginal revenue is equal in both markets and in turn equal to marginal cost. Output and its distribution is of course fixed conditional on the degree of restriction that is determined exogeneously by the authorities.

Suppose that the degree of restriction is raised by the authorities, so that the domestic demand curve is shifted to the left. Total demand is thus reduced at the ruling price, and the firm is faced with the possibility of reducing export price in an effort to maintain its profit position. If marginal costs are rising in the neighbourhood of the initial output before the rise in the degree of restriction, the reduction in home demand will reduce the marginal cost of producing output for the foreign market below its marginal revenue, hence encouraging an expansion in sales to that market. In general the price charged to the home market will rise and sales fall, while the price charged abroad will

fall and sales increase. On the other hand if marginal costs are falling in the neighbourhood of the original equilibrium output, the effect of the reduction in home demand will be to raise the marginal cost of output above marginal revenue received in the foreign market, hence leading to a reduction in the volume of output sold abroad and a rise in its price.[1] If marginal costs are constant the volume exported will remain unchanged. It follows that the reaction of the volume exported in the context of the price discrimination model to a rise in the degree of restriction in the home market depends critically on the nature of the cost curves in the region of the initial equilibrium output. In any given case the outcome is uniquely determined by the facts with regard to costs. If diminishing returns are in operation, marginal costs will be rising, but this is of course not necessary for a maximum profit position to exist. Marginal costs may be falling but marginal revenue must be falling faster for an equilibrium to be possible in this region.

It may be argued however that price discrimination and full profit maximisation in the short run does not characterise the behaviour of British producers in export markets. If this is the case an alternative model may be suggested.

A possible hypothesis is that prices are "fixed" (subject to the level of money wages) and firms act as quantity takers. Business men are not "marginalists" in this case but adopt some rule-of-thumb pricing or follow what is regarded as the "normal" price. In this case output produced will equal quantity demanded, and since the price of the product *ex hypothesi* remains fixed no reduction in home demand will lead to any increase in the volume supplied abroad unless excess demand exists abroad. This is a necessary, but not a sufficient condition, for such an increase.

There is however a further consideration. Even if in the short period the firm operates under conditions of fixed prices, there will exist a limit to the volume of output supplied, insofar as the firm realises that beyond a critical point marginal costs will exceed price (unless constant or decreasing costs prevail throughout) and some units will be sold at a loss.

The fixed price model does away with full profit maximisation but not necessarily with price discrimination. Fixed prices may

[1] For a more rigorous statement of the argument see the Appendix below.

be the rule in both markets over the relevant period but they may not be equal. Thus a rudimentary realisation of elasticity differences between markets may enter into price policies.[1]

The fixed price model is complicated in its simplest form by the fact that once prices are fixed and the policy is to sell what can be sold at these prices, there is no mechanism for allocating output between the two markets when excess demand exists in both. This is clear if it is assumed that the same price is charged in each market. Determinacy can be injected into the problem by the assumption that the domestic price is likely to be equal to or greater than the foreign price, and that there is a tendency for the domestic market to be satisfied, exports becoming a residual. A less stringent assumption is that a minimum level of exports is maintained or a minimum proportion of total output is exported, demand conditions permitting. A policy of this kind may be pursued either because it is desired to maintain goodwill in the foreign market or perhaps from the fear that if exports fall below that minimum the firms may incur such official displeasure that uncongenial policies restricting the home market may follow.

Given a minimum level of exports the firm will sell as much as it can in the domestic market at the ruling price. Sales to the home market will not exceed the point at which further output produced for sale at home is sold at a loss, which is conditional on the minimum export level. Given the level of home demand at the ruling price and level of restriction, the amount exported after satisfying both home demand and the minimum export level will depend on the relation of price charged in the foreign market to the marginal cost of production and the level of foreign demand at that price. The peculiarity of this case is that a reduction in home demand following a rise in the degree of restriction will only result in an increase in the level of exports if there is excess demand abroad at the ruling price charged abroad and the previous level of restriction. Excess demand abroad is a *necessary* but not a *sufficient* condition that exports will increase if the degree of restriction at home is raised, for excess demand at home

[1]The notion of "full cost" or "just" pricing by manufacturing industry is often criticised on the ground that it cannot explain price discrimination. This criticism rests on the assumption that the profit margin or mark-up over cost is fixed by the firm entirely independently of any suspicions about demand, but there is really no reason why this should be so.

may have existed at the old level of restriction and the rise in restriction may only serve to reduce excess demand at home rather than the volume of sales.

Suppose now that there exists the possibility of shifting the foreign demand curve by introducing "selling costs".[1] Selling costs are defined for present purposes to include not only expenditure on advertising but more important still expenditure on services associated with marketing and maintaining the product. In the case of durable goods shipped abroad, foreign choice depends very much on the services provided in parallel with purchases such as maintenance facilities and an adequate flow of replacements and spare parts.

Assuming that in the price discrimination model an increase in the volume of "selling services" will shift the foreign demand curve to the right, but at a diminishing rate, a rise in the degree of restriction will increase the volume of exports supplied provided that marginal costs are rising and that the initial position is stable.[2] If marginal costs are falling the reverse is true provided again that the initial position is stable, while constant marginal costs leaves the volume of exports unchanged.

In the fixed price model, with a minimum level of exports the firm determines the maximum level of production for export, given this minimum level and the level of home demand. Since foreign demand is now no longer fixed but can be varied by varying the volume of selling services, the firm will now produce for export up to the point where the sum of marginal production cost and selling cost is equal to the price charged in the foreign market. Of course the level of home demand which by assumption is supplied first may be such that this equality cannot be realised without violating the restriction imposed by the minimum export level. A reduction in the level of home demand can lead to an increase in the volume of exports provided the existing export level is above or equal to this minimum level without excess demand being present in the home market.[3] The introduction of selling services removes the necessity for excess demand abroad to

[1] To simplify the problem it is assumed that selling costs are not involved in the home market.

[2] The proof of this proposition is given in the Appendix below.

[3] Again the behaviour of marginal costs in the neighbourhood of the initial position is important. See the Appendix below.

exist in order that exports shall be substituted for domestic sales, for the possibility of variation in selling costs introduces an adjustment mechanism between demand and supply.

II

Three main factors emerge from this brief analysis as important in the debate.

(1) The market structure that exists.

(2) Cost conditions in the neighbourhood of current working.

(3) The extent to which the demand for exports can be increased by other than price changes, i.e. by an expansion of selling services.

With reference to the existing market structure it is probably reasonable to suppose that British industry tends to be imperfectly competitive in a broad sense, in so far as it is faced with negatively shaped demand curves both at home and abroad. There is clearly also a presumption that foreign demand is likely to be more elastic at ruling prices than home demand since the possibilities of substitution given the strength of Britain's leading competitors are probably greater. Under these conditions full profit maximisation would dictate the practice of price discrimination.

The question is (*a*) whether firms practise price discrimination, (*b*) whether they maximise short run profits in the manner required by the model.

It is hardly possible to reach any clear cut conclusions about the practice of discrimination. One might suppose that in view of the route taken by international views on commercial policy since the war, discrimination on a wide scale would meet with official disapproval. Slips from grace have undoubtedly occurred, such as the alteration in the sterling price of certain commodities after the devaluation in 1949, about which nothing was done by the authorities, but by and large one would suppose that memories of the bitter inter-war price struggles and the prevalence of dumping would serve to maintain a certain degree of "fair practice" in international trading.

This does not ensure of course that price discrimination does not occur, indeed there appears to be some evidence that it does in

certain industries, although it is hard to generalise.[1] Even if discrimination is not overt, concealed discrimination may take place by providing better quality goods or extra services and accommodation to foreigners at the price that rules in the domestic market. It is therefore not possible to rule out the conclusions of the price discrimination model on the ground that discrimination is not practised, and so the model is not applicable.

A far more important reason for doubting the application of the model is the argument that short run profit maximisation is unlikely to take place. This raises a whole set of queries about the positive value of the kind of approach to the theory of the firm that the discrimination model exemplifies, and is too broad a subject to be given extended treatment here. It is sufficient however to emphasise that there are many reasons why price will not be adjusted continually to output or demand changes. If it is assumed that price stability is a prerequisite of sound long term planning and that short period fluctuations in output are regarded as temporary, price will tend to be adjusted to the long run "normal" level rather than altered in response to short run output changes.

It is arguable that industrial prices in the United Kingdom are likely to be sensitive to changes in factor prices but fairly insensitive directly to changes in demand. If that is so, the assumption of a constant price level conditional on the level of factor prices seems fairly reasonable. In this case one arrives at some version of the fixed-price model as the relevant tool of analysis.

If the discrimination model were accepted, the outcome of the argument rests on the nature of costs in the neighbourhood of the initial position. In a rather imprecise way, industrialists seem to have espoused some form of the "cost" argument contending that a reduction in output will only lead to higher costs and prices. However the analysis of Section I turns on the behaviour of short run marginal costs and it is hardly plausible that these were falling in the boom periods that credit policies have been designed

[1]It should perhaps be made clear that discrimination is not necessarily implied by different final prices—obviously in so far as costs differ between markets (including costs of transportation). Price discrimination involves establishing price differentials for an identical product at the factory gate.

to frighten away. With rising marginal costs the price discrimination model would suggest that credit restrictions would make it profitable to cut prices to the foreign buyer.

Rejecting this model leaves us with the fixed price model for which (in the absence of selling services) excess demand abroad must exist for the restrictive policy to lead to a rise in exports. There has in fact been little evidence that excess demand for British products has existed abroad over the last five years, and the concentration on supply limitations implied by the fixed price model in general seems inappropriate.

It seems clear that the problem of increasing the export volume is not one of production but of selling. This suggests that the main argument in favour of the direct effort of credit restriction on exports is that it will induce an expansion in the volume of "selling services" so that the foreign demand curve will be shifted to the right.

The main objection to this argument is that the nature of such "selling services" is usually such that an expansion of the type envisaged requires both a considerable period of time before they can be completed and a fairly substantial investment.[1] Closer appreciation of foreign needs may demand redesign of the product and durable goods exports will require heavy investment in service facilities which are still very much lacking in overseas markets. Short period credit restriction is unlikely to encourage expansion of this kind which must be based on longer run considerations of profitability. Furthermore a buoyant home market contributing to a high over-all rate of profit on capital is likely, as industry has argued, to encourage such investment and provide resources for its realisation. The creation of excess capacity will hardly encourage the introduction of new equipment and expanded investment abroad of the type discussed.

It seems fair to conclude that the argument that restricting credit will have a direct effect on the expansion of exports in the

[1]Those who doubt this should review the campaign conducted by the Renault company to expand their share of the U.S. car market. Mr. Balogh long ago diagnosed the export problem in terms of the need for investment. "The failure of British (and European) exports to the U.S. to expand was not mainly due to higher prices. Risk and the need of preliminary huge investment in advertising and merchandising played a more important role".
T. Balogh, *The Dollar Crisis* (Basil Blackwell 1949) p. XXX.

short run is highly dubious. In the short run, the level of overseas demand is a big problem, and it must be accepted that weakness in the face of autonomous changes in foreign demand is a characteristic of the British economy at the present time[1]. For the longer run this difficulty must be faced, but short-term restrictive credit policies hardly provide a starting point. This is not to say that credit restriction may not be justified on other grounds, related to the reduction in imports, the restraint of rising wages and prices or the effects of higher interest rates on the short term capital flow. In general these aspects of credit restriction provide a more plausible justification of tight money (although they too can be debated) than arguments based on direct effects on exports that have been briefly examined in this paper.

R. J. Ball

Manchester

Appendix to Section II

Define the following variables : R_h = revenue from domestic sales, R_f = revenue from exports Π = profits, C = production costs, X_h = volume of domestic sales, X_f = volume of exports, r = degree of restriction imposed by the authorities. Then in the price discrimination model we have

$$R_h = R_h(X_h, r) \quad \frac{\partial R_h}{\partial X_h} > 0 \quad \frac{\partial^2 R_h}{\partial X_h^2} < 0 \quad \frac{\partial R_h}{\partial r} < 0 \quad \frac{\partial R_h}{\partial X_h \partial r} < 0 \qquad (1)$$

$$R_f = R_f(X_f) \quad \frac{\partial R_f}{\partial X_f} > 0 \quad \frac{\partial^2 R_f}{\partial X_f^2} < 0 \qquad (2)$$

$$C = C(X_f + X_h) \quad \frac{\partial C}{\partial (X_f + X_h)} > 0 \quad \frac{\partial^2 C}{\partial (X_f + X_h)^2} \lessgtr 0 \qquad (3)$$

so that profit is defined by

$$\Pi = R_h(X_h, r) + R_f(X_f) - C(X_f + X_h) \qquad (4)$$

First order conditions for maximum profits require that

$$\frac{\partial R_h}{\partial X_h} - \frac{\partial C}{\partial (X_f + X_h)} = 0 \qquad (5)$$

$$\frac{\partial R_f}{\partial X_f} - \frac{\partial C}{\partial (X_f + X_h)} = 0 \qquad (6)$$

[1]For example the sudden shift in the U.S. demand for foreign cars occasioned by the success of the new "compacts".

Differentiating (5) and (6) totally with respect to r and solving for $\frac{dX_f}{dr}$

$$\frac{dX_f}{dr} = \frac{\begin{vmatrix} \frac{\partial^2 R_h}{\partial X_h^2} - \frac{\partial^2 C}{\partial (X_f + X_h)^2} & -\frac{\partial^2 R_h}{\partial X_h \partial r} \\ -\frac{\partial^2 C}{\partial (X_f + X_h)^2} & 0 \end{vmatrix}}{\begin{vmatrix} \frac{\partial^2 R_h}{\partial X_h^2} - \frac{\partial^2 C}{\partial (X_f + X_h)^2} & -\frac{\partial^2 C}{\partial (X_f + X_h)^2} \\ -\frac{\partial^2 C}{\partial (X_f + X_h)^2} & \frac{\partial^2 R_f}{\partial X_f^2} - \frac{\partial^2 C}{\partial (X_f + X_h)^2} \end{vmatrix}} \tag{7}$$

Stability requires that the principle minors of the denominator of (7) alternative from negative to positive. If marginal costs are rising $\frac{\partial^2 C}{\partial (X_f + X_h)^2} > 0$ so the denominator and numerator are both positive. Stability conditions are still satisfied if $\frac{\partial^2 C}{\partial (X_f + X_h)^2} < 0$ and $\frac{\partial^2 R_h}{\partial X_h^2} - \frac{\partial^2 C}{\partial (X_f + X_h)^2} < 0$. In this case $\frac{dX_f}{dr} < 0$ so that exports will increase if r increases and marginal costs are rising and will fall otherwise (except in the case of constant costs when $\frac{\partial^2 C}{\partial (X_f + X_h)^2} = 0$).

Introducing $S =$ selling services into (2), we have

$$R_f = R_f (X_f, S) \quad \frac{\partial R_f}{\partial S} > 0 \quad \frac{\partial^2 R_f}{\partial S^2} < 0 \tag{8}$$

and letting v denote the price of selling services, profit is now

$$\Pi = R_h (X_h, r) + R_f (X_f, S) - C (X_f + X_h) - vS \tag{9}$$

First order conditions for maximum profits require that

$$\frac{\partial R_h}{\partial X_h} - \frac{\partial C}{\partial (X_f + X_h)} = 0 \tag{10}$$

$$\frac{\partial R_f}{\partial X_f} - \frac{\partial C}{\partial (X_f + X_h)} = 0 \tag{11}$$

$$\frac{\partial R_f}{\partial S} - v = 0 \tag{12}$$

Differentiating these conditions totally with respect to r and solving for dX_f/dr, we have

$$\frac{dX_f}{dr} = \frac{\begin{vmatrix} \frac{\partial^2 R_h}{\partial X_h^2} - \frac{\partial^2 C}{\partial (X_f+X_h)^2} & -\frac{\partial^2 R_f}{\partial X_f \partial r} & 0 \\ -\frac{\partial^2 C}{\partial (X_f+X_h)^2} & 0 & \frac{\partial^2 R_f}{\partial X_f \partial S} \\ 0 & 0 & \frac{\partial^2 R_f}{\partial S^2} \end{vmatrix}}{\begin{vmatrix} \frac{\partial^2 R_h}{\partial X_h^2} - \frac{\partial^2 C}{\partial (X_f+X_h)^2} & -\frac{\partial^2 C}{\partial (X_f+X_h)^2} & 0 \\ -\frac{\partial^2 C}{\partial (X_f+X_h)^2} & \frac{\partial^2 R_f}{\partial X_f^2} - \frac{\partial^2 C}{\partial (X_f+X_h)^2} & \frac{\partial^2 R_f}{\partial X_f \partial S} \\ 0 & \frac{\partial^2 R_f}{\partial X_f \partial S} & \frac{\partial^2 R_f}{\partial S^2} \end{vmatrix}} \qquad (13)$$

Stability requires that the denominator of (13) be negative which will be the case if $\frac{\partial^2 C}{\partial (X_f+X_h)^2} > 0$, $\frac{\partial^2 R_f}{\partial X_f^2} \frac{\partial^2 X_f}{\partial S^2} - \left(\frac{\partial^2 R_f}{\partial X_f \partial S}\right)^2 > 0$ and all other signs are as given. In this event the numerator is also negative so that $\frac{dX_f}{dr} > 0$, i.e. a rise in the degree of restriction will raise the volume of exports. Provided the stability conditions are fulfilled when $\frac{\partial^2 C}{\partial (X_f+X_h)^2} < 0$ the numerator of (13) will be positive and a rise in the degree of restriction will reduce the export volume.

In the case of the fixed price model, define $\overline{X}_f =$ the minimum export level : $X_f^* =$ the maximum amount that will be supplied to foreigners. Given the level of home demand X_h, the firm will examine the relation between marginal cost and export price, (P_f)

$$P_f \gtreqless \frac{\partial C}{\partial (X_h + \overline{X}_f)} \qquad (14)$$

If the inequality is positive the firm will be prepared to produce $X_f^* > \overline{X}_f$ for export, and if negative it will only be willing to export $\overline{X}_f$. If actual foreign sales X_f is such that $\overline{X}_f < X_f < X_f^*$ a reduction in home demand will only increase X_f^* and not X_f. However if $X_f \geq \overline{X}_f$ such that (14) is an equality, the presence of excess demand abroad implies that

$$\frac{dX_f}{dX_h} = -1 \tag{15}$$

i.e. a reduction in homes sales will lead to an expansion in the export volume.

Introducing selling services with fixed prices, foreign demand can be written

$$X_f = X_f(S) \quad \frac{\partial X_f}{\partial S} > 0 \text{ and } S = X_f^{-1}(X_f) \text{ (say)} \tag{16}$$

so that "desired" exports will satisfy

$$P_f - \frac{\partial C}{\partial (X_f + X_h)} - v \frac{\partial X_f^{-1}}{\partial X_f} = 0 \tag{17}$$

where v is the price of selling services as before. However the level of X_h may be such that $X_f < \overline{X}_f$ in order to satisfy (17), and in this case a reduction in the level of X_h will not in general lead to a rise in exports. If however $X_f > \overline{X}_f$ and (17) is satisfied a reduction in home demand may lead to a rise in exports.

Taking the total derivative of (17) with respect to X_h, we have

$$\frac{dX_f}{dX_h} = \frac{\dfrac{\partial^2 C}{\partial (X_f + X_h)^2}}{-\dfrac{\partial^2 C}{\partial (X_f + X_h)^2} - v \dfrac{\partial^2 X_f^{-1}}{\partial X_f^2}} \tag{18}$$

which is negative provided marginal costs are rising and there are diminishing returns to selling services. If marginal costs are falling either outcome is possible depending on the extent to which diminishing returns to selling services outweigh diminishing marginal costs.

Restrictive Practices and After

1. The Study Outlined

The 1956 Restrictive Trade Practices Act has resulted in the registration of over 2,300 restrictive agreements, and about one-half of them have since been abandoned. What has happened as a result of these cancellations? Businessmen have frequently foreseen gloomy prospects: are their fears justified? Economists emphasize the advantages of competition: are their theoretical expectations confirmed? Public policy makers should know the consequences; the next step must depend partly on the outcome of the present legislation.

The study does not provide all the answers. It is a contribution which it is hoped others will continue and amplify. For one thing, the long-term consequences are of most importance, and it is too soon to discover them. Nevertheless, short-run changes are interesting in themselves, they provide clues to the future, and point to the significant industries which subsequently could be examined in more detail. Moreover, as a result of earlier analysis,[1] price agreements only were studied, since these were the most prevalent. (Price was taken to include "discounts" and "margins"). Only one of the agreements studied also involved production quotas. Finally, the consequences of abandoning registered agreements that were cancelled after mid-1959, and of agreements that were abandoned before registration, were not examined. Mid-1959 was the closing date to allow the elapse of at least one year before conducting the enquiry; this was to give time for the modification of former codes of business behaviour.

Requests for information lacked the authority that, for example, the Monopolies Commission could command. Individual persuasion was reinforced by the moral support of the Federation of British Industries and the National Union of Manufacturers, to whom I am extremely grateful. (They are not, of course, in any way responsible for or committed to our findings). But I am most indebted to the 262 firms and trade associations which provided detailed replies. I wish, however, that the 500 or so

[1]See my article "Freer Prices—What Progress?" *The Banker*, February 1960.

others from which enquiries were made had been more co-operative. In several respects, therefore, this enquiry cannot claim to be complete or comprehensive.

It was felt essential to obtain information by personal enquiry.[1] Questionnaire forms, identical with those issued to the 61 interviewers which provided the framework for their reports, were also sent to a large number of firms and trade associations. Thus the basic information was standardized, with varying amounts of supplementary data.

In addition to the general enquiry, more detailed information was obtained from the electric cables and constructional steelwork industries. It had been my intention to conduct a greater number of detailed enquiries, to give them greater weight, and to co-ordinate the study through the wider survey. Unfortunately this emphasis proved impossible to achieve.

In what ways are the results biased? First, a large number of firms and trade associations did not co-operate. From 734 enquiries there were 220 detailed responses in standard form, and 42 less other detailed replies. A further 160 replies declined assistance; these when classified, were not entirely valueless. I assume that on balance those who failed to reply were less affected by the abandonment of the restriction than those who responded. Firms which feel they have been damaged or assisted by the 1956 Act are likely to want the fact known. Some firms ceased to exist because of free pricing; liquidators replied to 7 requests for information, and there may have been others who failed to reply, but the number involved is unlikely to be large. On balance there is probably a bias in the detailed replies in favour of much rather than little change.

With respect to some agreements not one firm would supply information. Out of a total of 240 agreements about which information was sought, details were obtained concerning 159; a further 14 abandoned agreements, re-registered in another form, were still operative; and the commodities subject to 4 other agreements were no longer in production—these trades had

[1] I am extremely grateful to 61 members of the Economics Association, each of whom conducted one or more interviews on our behalf, and to Mr. G. Roberts, who did much of the "office" work and who also conducted a number of interviews, particularly in connection with the cables industry (see below).

ceased to exist. Since the median number of members per agreement in the non-response group was 6, and in the response group was 15, the agreements for which detailed information was unobtainable were in the more oligopolistic trades, in which relatively little short run change may be expected. These two overlapping biases both exaggerate the apparent overall effect.

The study is intended to examine the consequences of the abandonment of agreements that took place before mid-1959. It would be unwise to infer from this the consequences of abandoning *all* registered price agreements, since one further bias would be present: those agreements abandoned first are likely to involve the least disadvantage to the parties concerned compared with those which remain until later. Associations which believe they benefit from their arrangements are likely to wait longer before abandoning than those whose agreement are thought to be ineffective or useless.[1] This further bias is likely to result in the present enquiry under-estimating the disadvantageous effects on the parties, and the advantageous effects on the public, of abandoning all price restrictions.

What reliability can be attached to the answers of businessmen? They were, for example, being asked to compare the present situation with what would have happened if the agreement had continued. The question would be particularly difficult where substantial changes had occurred. In the 30 or so interviews which I conducted I was impressed by the care with which such answers were thought out. It cannot of course be assumed that all were reliable, however honest the attempt. Occasionally it was possible to check reliability, sometimes from published statistics, but more often from information provided by other firms in the same or related industries. (Businessmen may of course have been agreeing about the replies they would give).

Members of some agreements referred to their trade associations for a joint reply. This was very unsatisfactory because the association could supply only a limited amount of reliable data, and they tended to relate conventional views (such as the

[1]It is reasonable to assume that agreements cancelled because of internal dissension are as likely to have taken place before mid-1959 as after that date.

terrible things that would happen; persistent enquiries never revealed that these consequences had actually occurred). Nevertheless, in some trades very significant changes had taken place, for which there was ample independent confirmation.

2. The Principal Findings

Competition

The direct question was asked " do you think that *if the agreement had not been cancelled* there would now be: less competition; about the same degree of competition; more competition?" Tables 1 (*a*) and (*b*) classify the results. The row "All replies" relates to all detailed questionnaire forms on which that question was answered. "Agreement" and "Trade" summaries were also compiled.[1]

TABLE 1(*a*)
CHANGES IN COMPETITION

	Number of Entries	If the agreement had not been cancelled, there would now be:—		
		Less Competition	About the same Competition	More Competition
All replies ...	180	30% (7%)	69% (17%)	1% (76%)
By agreement ...	124	32% (16%)	67% (35%)	2% (47%)
By trade	113	33% (19%)	67% (40%)	1% (40%)

NOTE: Figures in parenthesis have been calculated on the "worst" assumption, that all non-responses would have replied "more competition."

[1]Agreements were grouped in trades on a product basis (only fragmentory information was available about markets).

Where opinions differed within an agreement or trade it was assumed that if at least one reply said "less competition" and no replies said "more competition", then the summary answer was "less competition". Where an equal number of firms in the same agreement or trade gave contradictory answers, as occurred with some answers recorded in Tables 2(*a*) and (*b*), the opinion of the largest firms carried most weight.

TABLE 1(*b*)
CHANGES IN COMPETITION

	Mining and quarrying	Non-metalliferous mining products	Chemicals	Metal manufacture	Engineering, etc.	Vehicles	Other metal goods
S.I.C. Order ...	2	3	4	5	6	7	8
Number of agreements	5	16	7	9	33	1	13
Less competition...	—	25%	43%	44%	48%	—	31%
About the same ...	100%	75%	43%	56%	52%	100%	69%
More competition	—	—	14%	—	—	—	—

	Textiles	Food, drink, tobacco	Wood and cork	Paper and printing	Other manufacturing	Distribution
S.I.C. Order	10	13	14	15	16	20
Number of agreements ...	8	8	9	5	5	5
Less competition ...	25%	37½%	11%	20%	—	20%
About the same	75%	50%	89%	80%	100%	80%
More competition ...	—	12½%	—	—	—	—

In 67 per cent of agreements there was believed to be about the same degree of competition as if the agreement had not been cancelled; and in 32 per cent there would now have been less competition. Since the latter figure reveals a positive response, it is the more significant; no doubt the 67 per cent included a number of "don't knows".

If the "worst assumption" is made about non-responses, that all would have been in the category "more competition",[1] this 32 per cent falls to 16 per cent, and it is suggested that the "true" figure lies between the two.

Why did so few businessmen believe that there would now have been more competition if the agreement had continued? One hypothesis being that the cancellation of the agreement would result in mergers which reduce competition below the former level. There may have been an unwillingness to admit that competition had been reduced. However, it will be shown later that so far few mergers appear to have been caused directly by the Act, and in only 7 or 8 agreements did it seem that concentration had significantly increased. Generally mergers were associated with "no change" in the degree of competition, or else conditions in the trade were still unstable: it is too early to say what the outcome would be.[2]

Table 1(*b*) distributes replies in Table 1(*a*) by agreement and by Standard Industrial Classification. Most response agreements came within Order 6, "engineering, shipbuilding and electrical goods" (some Orders have been omitted from the Table). The largest proportionate increases in competition were in Order 6, Order 5 "metal manufacture", and Order 4, "chemicals and allied trades".

Is the theoretical expectation confirmed that the response of "no change" would be found in the relatively oligopolistic agreements? The median number of members in agreements in which it was reported that "competition was about the same" was 10, while in agreements in which competition would now

[1]If the most plausible assumption is made, that all non-responses would have been in the "about the same" category, then of course the 32 per cent figure would still have fallen to 16 per cent.

[2]To illustrate, in one agreement of ten firms the two largest had merged, but neither would assist our enquiry. Interviews with two others revealed that the new combine was trying to discipline the trade by price cutting. One other firm in that industry replied that it was unable to help us because it was no longer producing the relevant commodities. In another year or so the situation in this industry might be transformed.

In another agreement three of the five members were partly-owned subsidiaries of the fourth member. Before the agreement was cancelled consolidation was effected, and the degree of competition was thought to be "about the same". These cases are not typical.

have been less but for the cancellation, the median was 15, thus confirming the expectation.[1]

It may be thought that many agreements were ineffective, which would account for the high "no change" percentage. Although most agreements were ineffective part of the time, it was reported that 72 per cent of the "no change" agreements were effective at the time of their cancellation, and that 80 per cent of all response agreements were effective.[1]

Was there any significant tendency for smaller rather than larger firms to report that competition had increased? Of 161 firms which supplied the relevant information, the median number of employees per firm was 230 in firms where competition had increased, and 200 where competition was about the same. Considering the range of firm size (from 15 to 3,718 employees in the former category, and from 6 to 11,460 in the latter), these figures are surprisingly close. In so far as the difference is significant, the explanation may be that the larger firm will typically produce a wider range of commodities within an agreement than the smaller, and will be more aware of changes in competition than some of its smaller more specialized rivals.

Prices

The first of several questions about prices was "what effect do you think the *cancellation of the agreement* had on *your prices* (or discounts—specify which) of the goods to which it related: are prices lower than they would otherwise have been; about the same; or higher than they would otherwise have been?" To emphasize that a comparison was required of present prices with what they would have been if the agreement had not been cancelled, the next question was "have your prices in fact changed?" Tables 2(*a*) and (*b*) show the replies, on the same basis as before.

[1]The median number of members in agreements in which reliable information was obtainable about competition was 13, compared with 15 for all response agreements.

TABLE 2(*a*)
CHANGES IN PRICES

	Number of Entries	What effect do you think the cancellation of the agreement had on your prices:—		
		lower than otherwise	About the same	higher than otherwise
All replies ...	215	29% (9%)	67% (20%)	4% (71%)
By agreement ...	146	32% (19%)	64% (39%)	5% (42%)
By trade ...	116	37% (23%)	59% (36%)	4% (41)%

NOTE: Figures in parenthesis have been calculated on the "worst" assumption that all non-responses would have replied "prices higher than they would otherwise have been".

TABLE 2(*b*)
CHANGES IN PRICES

	Mining and quarrying	Non-metalliferous mining products	Chemicals	Metal manufacture	Engineering, etc.	Vehicles	Other metal goods
S.I.C. Order ...	2	3	4	5	6	7	8
Number of agreements	5	17	6	9	33	1	13
Prices lower ...	—	41%	50%	44%	52%	—	31%
About the same ...	100%	53%	50%	56%	48%	100%	46%
Prices higher ...	—	6%	—	—	—	—	23%

	Textiles	Food, drink, tobacco	Wood and cork	Paper and printing	Other manufacturing	Distribution
S.I.C. Order	10	13	14	15	16	20
Number of agreements	8	25	11	5	5	8
Prices lower	—	16%	18%	20%	20%	37½%
About the same ...	87½%	80%	73%	80%	80%	62½%
Prices higher	12½%	4%	9%	—	—	—

In 32 per cent of agreements it was believed that prices were "lower" than they would otherwise have been, and in 64 per cent that prices were "about the same" (classified by "all replies" the latter figure was 67 per cent; of these, exactly 50 per cent said that prices had in fact changed, and so did 66 per cent of all replies). If the worst assumption is made about non-responses—that prices were higher than they would otherwise have been—then the figure of 32 per cent (classified by agreement) falls to 19 per cent, and the "true" figure may lie between the two.[1]

From Table 2(*b*), the largest proportionate price reductions, below the level that they would otherwise have been, are in Orders 6 and 4.

The expectation that agreements hold some prices down when demand is high receives weak support: in 5 per cent of agreements it was believed that prices were now higher than they would otherwise have been. In one half of these there was reported to be a strong sellers' market. The most prominent feature of the remainder was that principally the small firms in an agreement raised prices, their larger rivals reporting that their prices were "about the same."

Since relatively oligopolistic agreements were associated with "no change" in the degree of competition, does the same hold for prices? The median number of firms per agreement was 16 in agreements where prices were "about the same" as they would otherwise have been, 15 where prices were "lower", and 13 where prices were "higher" (the last figure is based on only 7 entries, and may be misleading). There is some slight indication that it was in the more oligopolistic agreements that prices rose.

Prices being "about the same" were reported more frequently from diversified firms than from specialists; while prices being "lower" were reported slightly more frequently from specialist

[1]It is desirable to have a method of checking at least a sample of the results. Published price indices are of value only for actual price changes (which were checked where possible); but there is no published data to show with any degree of reliability what prices might have been if the restriction had continued. The best method, which has its obvious limitations, is to obtain several opinions in a trade, preferably including some from non-members of the agreement.

than from diversified firms.[1] This would be consistent with a hypothesis that diversification increases liquidity and reduces the impact of market forces on an enterprise ; since 72 per cent of highly diversified firms reported prices "about the same" compared with 59 per cent of highly specialized firms, the hypothesis is further supported. It should be mentioned that diversified firms in this response category were larger than specialized firms, the median number of employees being 200 and 155 respectively, which may also be importantly relevant.

Moreover, because many firms which reported that prices were lower than if the agreement had continued were diversified, some may have been able to raise the prices of other commodities in compensation; the overall effect on a price index would therefore be uncertain.

Supplementary information showed that surplus capacity was found more often in agreements in which prices were "lower" than otherwise, although the reasons for surplus capacity varied widely. In a few agreements, there was unmistakable evidence that the agreement had enabled inefficient firms to survive,[2] and it appeared that the larger firms had sufficient capacity to absorb the greater demands on them. In electric cables (see below) a product and process innovation had caused surplus obsolete capacity, and the industry was in the process of adjustment; and in constructional steelwork (see below) a sharp temporary fall in demand was superimposed on a steady shift to a rival product.

It might be supposed that where non-members of an agreement sold a substantial proportion of the relevant products, then the cancellation of an agreement would make little difference; and that the more important non-members are, the smaller will

[1] A "diversified firm" is here defined as one in which less than 20 per cent of turnover is in goods subject to the agreement; and a "specialist firm" is one where more than 70 per cent of turnover is covered by the relevant agreement. Of 111 firms reporting prices "about the same", 44 per cent were diversified, and 24 per cent were specialized. Of the 51 firms reporting that prices were "lower", 31 per cent were diversified and 35 per cent were specialized.

[2] Evidence before the Restrictive Practices Court showed that this also applied to the Cotton Yarn Spinners' Industry. Unfortunately the government scheme so overshadowed the effects of the cancellation of the minimum price restriction that our enquiry into this industry, and into the Yarn Doublers' industry, had to be abandoned. These agreements were excluded from the study.

be the effects of cancellation. Table 3 shows that, except for a single agreement, the reverse appears to be true.

TABLE 3

RELATIONSHIP BETWEEN IMPORTANCE OF NON-MEMBERS AND PRICE CHANGE

Of the 31 agreements with no outsiders, 84% reported "prices about the same"

Of the 39 agreements with less than 10% sales from outsiders, 69% reported "prices about the same"

Of the 42 agreements with 10%–40% sales from outsiders, 55% reported "prices about the same"

Of the 6 agreements with 41%–60% sales from outsiders, 33% reported "prices about the same"

1 agreement with over 60% sales from outsiders, reported "prices about the same"

There is the possibility of a bias, but the trend is so marked as to make it unlikely. However, the number of members per agreement, in these four principal categories in which prices were reported "about the same", shows some tendency to increase with the percentage of sales done by outsiders: agreements with few outsiders producing the same products also have few members, and agreements with a high proportion of non-members have many members. The median number of members in the four principal categories are, in the same order: 9, 10, 21, 3. (The last figure is the median of only two observations, and may be misleading.)[1] Thus there is some tendency for the more oligopolistic agreements to be associated with a high response of prices "about the same," a relationship not found between the principal categories discussed above.[2]

What were the principal relationships between answers to the "competition" and the "prices" question? Of the 164 replies in which both questions were answered, 71 per cent of firms which reported increases in competition also reported that prices were lower than if the restriction had continued; and 83 per cent of firms in which competition was "about the same" also reported that prices were "about the same" (and 13 per cent reported that prices were "lower" than otherwise). Thus to a high degree

[1]These medians relate only to agreements in which prices were "about the same". The medians of *all* agreements in the five classes of importance of non-members are, in the same order: 9, 19, 13, 10, 10.

[2]See page 181, paragraph 4.

the effects of removing restrictions on the strength of competition appeared to be reflected in prices.

It was asked whether there is (or was) any evidence of a "price war". We did not attempt to define "price war", but left it to the interrogators to elaborate. In 35 per cent of agreements there was at least some evidence of price cutting, although this gives an exaggerated impression. A few firms appeared to regard the emergence of price competition as constituting a "price war"; their reports were not confirmed by rivals. In 10 agreements price cutting came *before* the abandonment, and may have caused it rather than the reverse. In five agreements it was reported that manufacturers were competing to offer higher discounts to distributors so that their products would not be neglected in favour of rivals offering better terms. In my judgment, on the basis supplementary information, *at most* a price or discount war could be said to have been caused by the abandonment of 18 agreements, that is 12 per cent of response agreements and 8 per cent of them all: even this is probably an exaggeration.

Asked "do you recognise a 'price leader' for the relevant goods?", 37 per cent of all replies said "yes", and 63 per cent said "no". Classified by agreement the figures were 41 per cent and 59 per cent respectively. In most cases it did not prove possible to identify the nature of the leader-follower relationship (but see the electric cable report below). Neither was it clear whether there had been price leadership within the agreement recently abandoned, or whether new leadership had arisen. The median number of firms per agreement where at least one firm reported that price leadership existed was 17 compared with a median of 15 for all response agreements.

Finally, enquiries were made about the exchange of price, cost and turnover information between members. Independently of this study, although arising out of it, trade associations and firms had been interviewed for the purpose of compiling a separate report.[1]

In 83 of the agreements concerned in the present enquiry, it was reported that arrangements had been made to exchange information. Most often this involved an exchange of prices

[1] "Some economic consequences of the Restrictive Practices Act", Part III of a Symposium. *Economic Journal*, September 1960.

either alone or in conjunction with costs and turnover. It will be seen from Table 4, which classifies 81 of them, that in 7 agreements the information related to turnover only; although in these cases the data were supposedly used primarily for collating industry statistics or for determining subscriptions to trade associations, in another trade, not included in this enquiry, it was found that an agreement to exchange turnover information had replaced a quota scheme for allocating output. Examples of the notification of prices and enquiries are given below, in the electric cables and constructional steelwork industries.

So far these agreements to exchange information have not been deemed registrable under the 1956 Act.

It is misleading to relate responses about price changes to the presence or absence of these agreements, since some were formed only *after* prices had fallen, as in the cables industry. Indeed, in those agreements in which price information is now exchanged a higher proportion reported that "prices were now lower than they would otherwise have been" than in all response agreements: 39 per cent against 32 per cent. This may suggest that price cutting led to agreements to exchange information, or that the agreements were ineffective in preventing prices from falling. The former explanation is the more likely since, to my knowledge, no information agreements were formed before 1959 (although some earlier ones accompanied restrictive agreements and were continued after the registrable parts had been cancelled), and it was not until March 1960 that the possibility of establishing them became widely known in Britain.

TABLE 4
INFORMATION AGREEMENTS

	P	PCT	PT	PC	C	CT	ENQ	T	Total
Number of agreements	40	17	7	1	2	4	3	7	81

NOTE: P=Prices; C=Costs; T=Turnover; ENQ=Enquiries

Mergers

The question was asked "has any firm acquired control of any other firm in your trade since October 1956 in which the

acquisition was, in your opinion, caused by the 1956 Act?" Would firms seek this alternative method of achieving the security afforded by a price agreement? Although in 24 out of 149 agreements for which replies were received it was stated that one or more mergers had been caused by the Act, supplementary information suggests that only in 15 agreements were the abandonments *closely* related to mergers, although in all the cancellation probably had some influence. In just over one-third of these agreements were also found price cutting, firms going out of business, or ceasing production of the relevant commodities, and competition having increased; in the remainder it was reported that competition had not changed. Almost all mergers were horizontal, although, as will be mentioned later, two agreements involving both horizontal and vertical mergers were part of the cables industry. Two other industries in which mergers took place have been mentioned already.[1]

It may be noted that the median number of members in these 15 agreements was 15, the same as the median for all response agreements.

It is important to know how the *concentration* of production has changed, and to know the availability of substitutes, for these will affect the strength and nature of competition. In 7 of these 15 agreements it seemed likely that market concentration had increased by an amount that could substantially affect competition in their respective trades. In 3 other agreements a few small and inefficient firms had been acquired by the more efficient. In some trades concentration was also increasing through firms ceasing to produce the relevant goods; however, since the information available from such firms was very limited, it could not be ascertained whether concentration was substantially affected, nor whether the Act was the principal causitive factor.

Although there is some support for the hypothesis that if the security a price agreement is said to provide is removed, firms will find the alternative of increased concentration through merger, it is certainly less than most commentators expected, and provides less than adequate confirmation for the hypothesis.

[1]See footnote 2, page 178.

There may be several explanations: the short time that has elapsed since the abandonment of these restrictions; demand has in general been buoyant; information agreements have been substituted; many trades are highly concentrated already in a few firms. The study points to one further reason, already mentioned briefly. The question was asked "how important were the goods subject to the agreement in the total turnover of your firm?" The following answers were received from 156 firms :—

40 per cent reported "less than 20 per cent of turnover"
21 per cent reported "20 per cent–45 per cent of turnover"
12 per cent reported "46 per cent–70 per cent of turnover"
27 per cent reported "over 70 per cent of turnover"

Clearly, a high proportion of firms were diversified. Also, as one would expect, the more diversified firms were larger than the more specialised; the median number of employees in the four categories were, in the order from diversified to specialist, 200, 225, 124, 115.

Although diversification does not prevent or reduce price competition, it may reduce the impact of unstable receipts, and therefore reduce the urgency with which a method of limiting competition is sought.

However, a firm may view the cancellation of an agreement rather differently if, although diversified, the other products which it produces are also subject to price regulation. At some stage these agreements may also have to be abandoned. The further question was asked: "If you were/are also parties to other registered agreements, how important were/are *all* goods subject to agreements in your turnover?" These were the replies:—

21 per cent reported "less than 20 per cent of turnover"
27 per cent reported "20 per cent–45 per cent of turnover"
14 per cent reported "46 per cent–70 per cent of turnover"
37 per cent reported "over 70 per cent of turnover"

Although, as one would expect, the average has shifted upwards, there is still much diversification into products not subject to restrictive agreements at all.

Cessation of production

The question was asked "has any firm in your trade (in the area covered by the agreement) ceased producing or selling the relevant commodities since the agreement was cancelled?"

Although in 19 per cent of response agreements it was reported that at least one firm was no longer in that market,[1] in most cases it proved impossible to discover whether the 1956 Act was a significant factor or not. Some were due to a combination of circumstances, and the cancellation of the agreement proved the "last straw". One firm in voluntary liquidation that I interviewed in the constructional steelwork industry illustrates the circumstances that may be present (see below).

Quality

In 10 agreements it was reported that quality had fallen. On three of them agreements about quality accompanied price restrictions, and in two of them these agreements continue ; in no case was quality reported to have fallen in matters covered by these agreements. In one of the ten, concerning mechanical equipment used in the building trade, evidently design had been cheapened. In another, also in the engineering trade, it was reported that after-sales service was less thorough and that research expenditure had fallen. In this case materials quality which was also subject to agreement, evidently had not changed. In another agreement, quality was maintained to the predominant buyer (a nationalised industry) but other purchasers were being given less elaborate equipment for the same price. After-sales service, which had been subject to agreement was evidently unchanged. In another case the principal producer, I.C.I., had not been a party to the agreement, but two smaller competitors, one specialised firm with 200 employees and the other more diversified with 150 employees, both reported that their quality had deteriorated. Small purchasers

[1]In addition, 28 firms failed to complete the questionnaire forms for the stated reason that they were no longer producing the relevant commodities, 7 firms reported that they had gone into liquidation, and 12 letters were returned by the Post Office marked "gone away". Of course there may have been some overlap between reports by firms still in the relevant trades, and those who evidently were not. The 35 firms that were obviously no longer producing the relevant commodities were spread amongst 14 agreements about which we had no other information, and 15 response agreements.

of this commodity probably would not be able to tell whether quality had fallen, but purchasers with laboratory facilities undoubtedly could judge. (For another example see the report below on the constructional steelwork industry).

Only in one agreement did it seem likely that a substantial fall in quality would constitute a public danger, although this level had not been reached. Of the seven members, two reported that their quality had fallen, and one replied that it had not. The large firm (1,600 employees) which was also highly diversified (in the sense used above), replied that quality had been maintained. The other two firms in which quality had deteriorated were both smaller (700 and 500 employees) and more specialised (46 per cent–70 per cent of turnover in this product in both cases).

Indeed, all firms except one which reported that quality had fallen were more specialised than the average.

It should be mentioned that a fall in quality is not necessarily undesirable. Some purchasers may wish to buy at a lower price even though quality is less good (in all the cases mentioned above prices had in fact fallen to below the level that they would otherwise have been). The restriction of price competition may lead to such emphasis on quality that consumers are offered better goods than are needed to satisfy the necessary functional requirements; price competition widens the opportunity for consumers' choice—providing that those who wish to buy the higher quality can still do so.

Efficiency

Where competition has increased there is additional stimulus to raise efficiency, but there are unlikely to be observable effects in the short period. Nevertheless, in five agreements, firms mentioned specifically to our interviewers that the cancellation of the restriction had led them to make vigorous attempts to become more efficient. Closer attention to costs and standardisation were mentioned in particular. On the other hand, two firms mentioned that, as a result of reduced incomes, research expenditure had been cut. Most firms when asked specifically about efficiency, in effect reported "no change". It is too soon to expect observable consequences.

Attitudes

The question was asked "at the time how did your firm regard the cancellation of the registered agreement: with anxiety; indifference; or satisfaction?" and "do you at present regard the outcome with: anxiety; indifference; or satisfaction?" The 159 responses are classified in Table 5. Little more than one-third of these firms were still anxious about the outcome of free pricing. The median size of the currently "anxious" firm was slightly below the average : 192 against 200 employees.

TABLE 5
ATTITUDES

159 responses from firms	At the cancellation of the agreement	At the time of the enquiry
Anxiety	50%	37%
Indifference	45%	53%
Satisfaction	5%	10%

3. The Detailed Studies

Before attempting to draw conclusions from the assembled replies, the results of the case studies will be summarised. These two industries were chosen because both exhibited clear signs of price cutting after the abandonment of the agreement, yet both were very different types of industry. I am particularly grateful to the firms and trade associations concerned for their co-operation. The opinions expressed are, of course, entirely my own.

Constructional Steelwork

There are 286 members of the British Constructional Steelwork Association.[1] About half the firms specialise on making and erecting steel frames for buildings, which is the principal product of the industry. Other types of self-supporting tower also come within the Association, but bridges do not.

The principal competing products are reinforced and prestressed concrete. However, steel is often preferred for flexibility in design or use, and steel frames are erected more quickly than

[1]This study was based on a 10 per cent sample of firms of all sizes, specialised and non-specialised in several localities.

their substitutes. Speed and reliability in erecting a frame are important features of competition, and often are more significant for winning orders than moderate differences in price.

The larger firms produce the entire range of customer requirements, and supply national and international markets. The smaller firms undertake more limited tasks, and specialise in local work. Typically firms operate in a localised oligopolistic market. Consumer loyalty to particular suppliers is variable; this reflects the range of work that is undertaken, and also the need for and the ability to offer speedy and reliable delivery. Transport costs are high in relation to the value of the product.

From 1936[1] to 1958 an agreement required the notification of enquiries to the Association, and specified maximum and minimum prices on a tonnage basis for each type of completed structure. The price schedules were extremely complicated and detailed. Shortly before the system was due to be investigated by the Restrictive Practices Court in January 1959, the price and tendering arrangements were cancelled voluntarily.[2]

The cancellation of the price agreement coincided with the first serious post-war recession in the industry. But by the time of this enquiry, October/November 1960, demand had substantially recovered, and many firms were reporting long order books. It also coincided with a change by many local authorities in the Fire Prevention Regulations. Steel frames for building were required to have thicker cladding; this gave an immediate advantage to reinforced concrete.

When the pricing arrangements were cancelled an agreement relating to the notification of enquiries was continued. When a firm receives an enquiry (for an order over 10 tons) the B.C.S.A. is notified. The reporting firm then learns from the Association the names of other firms tendering for the contract. The winning

[1]The price scheme was associated with the substantial changes in the steel industry's arrangements, which followed the imposition of tariffs in 1932 and the reports of the Import Duties Advisory Committee.

[2]The Court heard evidence on a number of agreements about standard forms of contract and arrangements with particular customers, but the Registrar invited the Board of Trade to authorise the removal of these agreements from the register (and therefore from the scope of the Court) as being, under section 12 of the Act, "of no substantial economic significance". This was done, and the case was closed.

firm notifies the Association, and rivals for that contract will be informed. Prices are not reported at any stage. Firms can therefore identify their rivals for a particular contract and can plan their competitive strategy accordingly. What consequences followed from the cancellation of the price agreement and from the continuance of this notification arrangement?

There are regional differences, but firms reported that prices fell between 10 per cent and 30 per cent, with the mean around 20 per cent. By the time of the enquiry prices had risen from these low levels, and most firms believed that prices were at or approaching the former minimum level; for some products, prices were still much lower than they would have been had the scheduled prices continued, and for others prices were already higher. Most firms believed that if demand continued to increase, prices would rise above the maximum level that would have been permitted under the former schedules. (No firm appeared to be making substantial use of the former price schedules, since it was not considered worth while to try to keep them up to date with factor price changes).

What consequences followed from price cutting? All firms reported that profits had fallen (which they would have done anyway, although probably not to the same extent; the short-term elasticity of demand seemed to be rather low). A few firms, mainly the more specialised, had postponed but not cancelled or otherwise amended investment plans. The important factor seemed to be uncertainty of the immediate future income (in relation to the reserves that would be needed to continue selling below total cost) rather than doubts about the level of future demand. Some firms reported that they were giving closer attention to costs than hitherto.

Research and development were virtually unaffected since research is conducted through the trade association not by individual firms. There is pressure to increase research expenditure from the industry as a whole.

Many firms reported that there was tendency for designs to be cheapened, the margin of safety between the actual design and the British Standard for the industry, B.S.449, being narrowed. But it was generally believed that some change in design would have occurred anyway, although not to such a wide extent.

Where new and more economical designs were developed, it seemed that the recession was in the main responsible rather than the cancellation of the pricing schedules. Several firms reported that some customers who bought very cheaply in the recession found the work unsatisfactory, and had subsequently to incur additional costs.

No mergers were reported. Because of the ease of entry and the individualism of managers, mergers were not considered useful or practicable.[1] Anyhow, most firms knew that the recession would be temporary, although there was uncertainty and anxiety about its likely duration. There is some diversification in this industry, and although a few firms were talking of further diversification it seemed to be a difficult step to take. Most of the diversification was into other forms of steel fabrication requiring different skills.

Almost every firm reported that the cost of tendering had increased since the B.C.S.A. system was abandoned. Formerly, unsuccessful tenderers received a small rebate from the successful firm to compensate for tendering expenses. But the rebate was a small proportion of the costs incurred, and its loss when the scheme was cancelled was trivial. Much more important was the fact that once the B.C.S.A. price schedules became inoperative, many purchasers, in particular the public authorities, invited tenders from more firms than hitherto. Simultaneously, because of the recession, firms were seeking enquiries from more customers. Thus the cost of tendering rose because firms were tendering for more contracts in relation to the number of orders received. A few firms resisted this tendency through shortage of design staff. Most firms now report that although demand has substantially recovered, they are still tendering for more work than hitherto.

As a result of the very low prices that ruled for more than a year, 4 works were closed. One firm in voluntary liquidation that I interviewed was closing down for a variety of reasons. It had specialised in producing a particular item of capital goods for export, and this market had vanished; the branch factory of a large motor firm was shortly to be erected nearby, and it could foresee the loss of its skilled workers attracted by high

[1]Where mergers took place in the past they were generally forward vertically from steel production, rather than horizontally.

wages. Some of its rivals also thought that its up-to-date equipment was being inefficiently utilised. The serious fall in structural steel prices was sufficient to send the firm into liquidation. Since there are many structural firms in the area, it did not seem as though consumers would suffer any disadvantage.

To what extent did productive efficiency count in winning orders in a recession? It was unimportant in most sections of the industry. For standardised and relatively simple work a large number of firms would be invited to tender, and the lowest price would win. I was told that the firm prepared to cut the lowest was frequently not the most efficient in production, although it may be acting efficiently in a more general sense. The "transfer mechanism", as used by Mr. Jack Downie,[1] in which the more efficient firm gains at the expense of the less efficient, appeared to have been interrupted by this recession.

What was the effect of the agreement to notify the receipt of an enquiry? In general, firms found this information useful because it enabled them to make more effective pricing decisions. A firm may discover the degree to which a rival's capacity is utilised, and can often estimate the production costs of its regular competitors. Knowing the names of its rivals for a particular contract enables a firm to make use of this information. In a recession, however, when many more firms tender than usual, and when the actions of rivals are less predictable, such information is of diminished importance. And some firms admitted that during the recession they failed to notify enquiries. Although knowing the names of rivals for a particular contract provides the opportunity to allocate contracts in a mutually convenient way, which could stabilise prices, I found no instance of this actually having occurred.

In conclusion, it seemed that although prices and profits remained very low for over a year, damage to the trade was negligible. But if recessions of this magnitude are repeated at all often, modernisation and investment plans are likely to be affected more seriously. The consequences of price cutting were not as severe as many firms feared, but perhaps this was because it followed many years of prosperity. On the other hand, customers paid lower prices and received a product that

[1] "The Competitive Process", Duckworth, 1958.

was usually, but not always, satisfactory for the purpose. When demand is high some consumers will have to pay more ; but this is not necessarily undesirable. Research and development were not adversely affected, and the development of new techniques, which would benefit consumers, was not hindered. The "shake-up" to the trade probably had beneficial effects in that greater attention was given to cost reduction.

Insulated Electric Cables

It will be misleading to consider only the electric cables industry, since the two earlier processes are closely related to it both technically and through ownership. In the first process, copper wire bars are purchased from overseas[1] and rolled into rods; 6 firms, members of the former Rod Rollers' Association, which was dissolved in October, 1957, dominate the industry. In the second stage, the 17 members of the High Conductivity Association and some non-members draw these rods into wire. Finally, the wire is made into cable by more than 30 firms, members of the Cable Makers' and similar associations, and non-members who at one time were grouped in the Independent Cable Makers' Association; most of these firms are now members of the newly-formed Electric Cable Makers' Federation.

The industry is highly integrated vertically. The principal rod rollers and some of the wire drawers are also engaged in later processes. Two principal fully integrated companies, British Insulated Callender's Cables Ltd., and Associated Electrical Industries Ltd., produce about 70 per cent of the output of cables and control different but high proportions of the earlier processes; one other company is fully integrated, three or four are integrated in two stages of production, and the remainder are non-integrated. At least one major company also has financial interests in the supply of copper. These integrated concerns send part of the output of intermediate products to their associated works, and sell the remainder.

The reports of the Monopolies and Restrictive Practices Commission provide a detailed survey of former practices in the industry. Price agreements and the loyalty and aggregate

[1]Two integrated companies also purchase the cheaper "blister" copper and "fire refine" it into copper bars wire and other products in this country.

rebates in wire drawing were found to be contrary to the public interest (wire rods were outside their terms of reference).[1] Their recommendations were not implemented because of the coming of the 1956 Restrictive Practices Act; however, the rebates were cancelled in 1957, and the price restrictions were abandoned voluntarily in June 1959. The High Conductivity Copper Association then adopted an information arrangement by which prices and, where relevant, the names of customers who receive special prices are exchanged voluntarily. One important partially-integrated producer and one smaller wire company have remained outside the system. As already mentioned, the Rod Rollers' Association was dissolved in October 1957.

Before the implementation of the Commission's recommendations in the cables industry,[2] all associations, except the Independents, operated common pricing arrangements (in some cases dating from the beginning of the century), and some associations regulated output by quotas, engaged in exclusive dealing, and operated three systems of rebates. After the Commission's report, substantially all that remained were minimum price agreements in rubber and thermoplastic cable (based on the lowest cost producer and supervised by the Board of Trade), mains cable, covered conductors, and high tension cables, together with individual quantity rebates. Except for high tension cable, these price arrangements were terminated voluntarily in May 1959.

Stimulated by a shortage of rubber during the war, by the early 1950's thermoplastic cable, for most purposes a cheaper and more efficient product, was offering severe competition to rubber-covered cable in the important general and house wiring sections of the industry. Pioneered principally by a non-member of the C.M.A., its popularity severely disturbed the members of the Association, whose capacity was mainly for the production of obsolescent rubber cable. One effect of this innovation was to lower barriers to new entry into this section of the industry (the new equipment cost about 10 per cent of the old, and was easier to operate), which had important consequences. Entry into the

[1] "The Supply and Export of certain Semi-Manufactures of Copper and Copper-Based Alloys", July 1955.

[2] "The Supply of Insulated Electric Wires and Cables," June 1952.

mains cable section of the industry was evidently less easy because of difficulties in the supply of suitable high-speed machinery. There are substantial economies of scale in rod rolling, but not in either cable making or wire drawing; however, there are economies of long runs. Consumer loyalty, an obstacle to new competitors in many industries, is stronger between the buyers and sellers of copper wire than in the general and house wiring section of the industry, where wholesalers are important. Difficulties for new competitors wishing to enter the industry are least in the most important section of the cables industry (as measured by turnover), higher in most other sections of the composite industry, and highest in rod rolling because of the large plant size in relation to demand that is needed for economical production.

These are the relevant basic facts. What consequences followed from cancelling these agreements?

Prices for rods have remained unchanged since before the cancellation of the Rod Rollers' Association. (The "prices" of these products are quoted on a "per ton" rate above the price of copper wire bars at the time the rods are purchased. Thus the considerable fluctuations in copper prices are shifted forward). This price stability can be attributed largely to the conscious inter-dependence between the few firms, re-inforced by vertical integration, and is all the more remarkable because some very high productivity rolling equipment has been installed in recent years. Likewise in the wire trade, although there was uneasiness because of events in the cables industry, the substantial degree of vertical integration together with the information reporting scheme have helped to maintain price stability; wire prices have, however, weakened slightly since the agreements were abandoned.

With the new thermoplastic cable becoming important, agreement on a realistic revision of prices for rubber covered cables was difficult, manufacturers were reluctant to scrap the obsolescent equipment, and the diffusion of this innovation throughout the industry was clearly delayed. Aggravated by a fall in both home and export demand, towards the end of 1958 there were signs of concealed price undercutting (partly for this reason one company left the Association, but rejoined later). And when the cable price agreements were cancelled in May 1959

prices fell by some 20 per cent. Rivalry between the major firms resulted in price cutting, which would probably have occurred anyway, although less dramatically, and others were forced to follow. In November, 1959, B.I.C.C., which had meanwhile consolidated its position as the leading firm in the industry, raised its prices by 7½ per cent for low voltage mains cables (in which there were 11 important competitors), and in January 1960 by 5 per cent for the more important general and house wiring cables (in which there were about 30 competitors). In both cases rival firms followed. Through its dominant position, and with skilful timing, B.I.C.C. established a position of price leadership much stronger than had existed before, and with a substantially lower level of prices.

After price leadership had been established, price information agreements were started also in the mains cable, general wiring, and covered conductor sections of the industry. And many firms that were outside the former C.M.A. system have now been brought into the Electric Cable Makers' Federation.

Structural change, which had been proceeding spasmodically since before the war, was accelerated after the cancellation of the quota scheme. Under the pre-1952 scheme it had evidently proved advantageous for A.E.I. to have its subsidiaries listed as separate companies; the abolition of the quota system was at least one factor in leading to the rationalisation both of production and research.

More significant, however, were the mergers that took place. They began shortly before prices fell, and continued until at least seven firms had been absorbed. In at least three companies rationalisation and the reduction of surplus capacity has resulted. No independent concern has gone out of business. These mergers had many motives.

Underlying some was the introduction by two companies of modern automatic equipment for rod rolling, mentioned above, which increased capacity and offered increased economies of scale. These firms were aware of the necessity to secure outlets. One firm, fearing for its market, asked a major company to take it over (which it did). Although the need to reduce obsolete capacity was a basic motive for merger, it gave rise to other mergers. For example, one firm lost an important customer by

merger to a rival, and it responded by securing an outlet for itself by merger. Another take-over was motivated partly through the regional buying habits of a publicly-owned corporation. One firm wished to enter a new market in the cables field, and it could conveniently do so by merger.

There is no doubt that the industry is more highly concentrated than before; but entry into the thermoplastic cable field is now easier, and through this substantial part of the trade new competition may in principle be introduced to other sections. Nevertheless, non-integrated cable makers may find themselves in an invidious position; their supplies have to be purchased through intermediate producers which are mostly vertically integrated, and which may also be cable makers. The several non-integrated producers of wire and one important producer who is integrated in the first two stages of production have reacted to the power of the fully-integrated producers (with what success remains to be seen); for example, some wire drawers have cut their margins, they report, to assist their non-integrated cable customers to remain in business. Even if the lower costs of the fully integrated producers justify lower final prices, the squeeze on non-integrated companies may or may not be the straightforward operation of the market resources; much depends on the way intermediate products sold outside the company are priced by integrated concerns.

Our enquiries did not produce evidence of quality deterioration during this period of low prices. Several firms mentioned that there was now less money available for research and development, but they did not seem to be reducing this expenditure. Most firms regarded price-cutting as transitory. Some firms were postponing investment in the industry until its future became clearer and its prices more profitable. There was evidence of diversification, both old and new, into products unconnected with cables.

It is difficult to know to what extent these changes were *caused* by the cancellation of the restrictions and the 1956 Act, by the Monopolies Commission, or whether they would have happened in any event. Certainly their *timing* was affected. The present structure of the industry is very highly concentrated vertically and horizontally, firms exchange price information,

and there is strong price leadership for the final products where there are many competitors. Price leadership and high concentration are not necessarily undesirable in themselves. A limited safeguard for the public interest is the apparently greater ease of entry for new competitors in the important general and house wiring cables, which accounts for about one-half of the turnover in cables, but to some extent the entry cost conditions are determined by the vertically integrated firms. Also there probably remain opportunities for technical change, which may provide a useful element of competition. Whether these safeguards prove to be at all adequate remains to be seen. The adequacy cannot be judged by casual observation. The structure of the industry, the behaviour of the dominant firms towards their rivals, and the performance characteristics would all need to be studied closely.

4. Conclusions

The short-term consequences of cancelling restrictive price agreements do not seem as widespread or as deep as many people expected. However, the magnitude of these changes depends in part on the economic climate and on special circumstances at the time of the cancellation and of the enquiry. With demand generally high, one would not expect the degree of competition between sellers to increase significantly. In less than one-third of agreements it was thought competition had increased because of the removal of the restrictions, and in almost all the remainder that competition was "about the same". Agreements with many rather than few members, and firms with rather larger than average employment experienced most increases in competition.

A high proportion of firms which reported increases in competition also reported that prices were lower than if the restrictions had continued, and a low proportion of firms in which competition was "about the same" reported that prices had fallen. Where demand is high and competition was not increased by cancelling the restriction, one would not expect to find price cutting; indeed, except where firms prefer long order books, one would expect some prices to rise. In a very small minority of agreements prices had risen. However, if the removal of restrictive agreements affects prices at all in the short run, and in

nearly two-thirds of agreements there was "no change", prices are more likely to be lowered than raised. Surplus capacity was more evident in agreements where prices were lower than where they were not, although the reasons for surplus capacity varied greatly.

Highly diversified firms reported significantly more frequently than specialised firms that prices were "about the same" as if the restrictions had not been cancelled; and, to a less marked extent, prices had been lowered in the specialised rather than the diversified firms. Specialist firms are also more likely to reduce quality under price competition than diversified firms. There was little difference in the number of members per agreement between agreements in which prices fell and agreements where prices were "about the same"; if anything, there was a tendency for price increases to be found in the more than in the less highly oligopolistic agreements.

"Price (or discount) wars" rarely took place when agreements were abandoned, although they were certainly not unknown.

Agreements to exchange information have become a common replacement for the former restrictive agreements. Most often the information concerns prices only, or prices, costs and turnover. Of restrictions abandoned before the early part of 1959, information agreements were probably established after rather than before prices had fallen (this is not so of agreements abandoned more recently, which were outside the present study). It is too soon to be judging the consequences for the stability and dispersion of prices.

Some mergers, principally horizontal but also vertical, have resulted directly from the abandonment of restrictions, yet to a smaller extent than was generally expected; concentration has increased in a few markets. Although the many mergers in recent years have substantially other causes, the Restrictive Practices Act may have been influential indirectly. To a surprising degree, firms formerly engaged in restrictive agreements were highly diversified; not only did this reduce the impact of increased competition where this occurred as a result of cancelled restrictions, but it also made it difficult to assess the net outcome of changes in the prices of those goods formerly subject to agreements.

Although many firms felt anxiety at the time of the cancellation of their restrictions, at the time of the enquiry (mid-1960) only about one-third of firms still felt some anxiety about the outcome; these firms were slightly smaller in size than the average.

It is difficult to judge how much consumers have gained from these changes. What is certain is that some consumers have benefited already, and may continue to benefit, principally through lower prices; in the long run greater stimulus to efficiency which increased competition, will bring further benefits. On the whole the disadvantages to the public in the short run seem to be rather small.

Readers must be reminded of the bias's in the data used for this study: the assumption about non-responses, and the examination of registered agreements abandoned no later than mid-1959. Also, it is essentially a study in the *short-term* consequences of cancelling a limited number of restrictions. It is desirable that someone with greater authority, such as the Monopolies Commission, should conduct a more widespread investigation, and should attempt to assess the longer-term consequences.

J. B. Heath

Manchester University,
Northwestern University.

Books Received

(Oct., 1960 — May, 1961)

BABEAU, Andre : *Les Conseils Ouvriers en Pologne.* Librairie Armand Colin. pp. 305.

BLYTH, C. A. : *The Use of Economic Statistics.* (Minerva Series, No. 5). George Allen & Unwin, London. 28/-.

BUSSCHAU, W. J. : *Gold and International Liquidity.* South African Institute of International Affairs, Johannesburg. pp. 99.

CARTER, C. F. : *The Science of Wealth.* An elementary textbook of Economics. (Publishers) Arnold. 14/-.

CORDEN, Max : *A Tax on Advertising.* The Fabian Society, London.

DESAI, M. B. : *Co-operation in Karjan.* Indian Society of Agricultural Economics, Bombay.

DISKALKAR, P. D. : *Resurvey of a Deccan Village Pimple Saudagar.* Indian Society of Agricultural Economics, Bombay.

GHOSH, Alak : *New Horizons in Planning.* The World Press Private, Ltd., Calcutta. 18/- net, pp. 212.

GOWDA, K. V. G. : *Appreciation of the Indian Rupee.* Chaitanya Publishing House. Rs. 15.00, 30/-, pp. 264.

GRAMPP, W. D. : *The Manchester School of Economics.* Stanford University Press, California. 25/-, pp. 138.

GULATI, I. S., *Resource Prospects of the Third Five-Year Plan.* Orient Longmans. Rs. 10.00, pp. 152.

HILTON, George W. : *The Truck System.* Including a History of the British Truck Acts. Heffer, Cambridge. 21/-, pp. 155.

HOSELITZ, Bert F., SPENGLER, Joseph J., LETICHE, J. M., McKINLEY, Erskine, BUTTRICK, John & BRUTON, Henry J. : *Theories of Economic Growth.* $7.50, pp. 334.

JASKARI, Osmo V. : *A Study in the Theory of Incidence of Taxation.* (Publishers) Suomalainen Tiedeakatemia.

JASKARI, Osmo V. : *Depreciation Allowances and The Real Capital.* Suomalainen Tiedeakatemia.

KIMBLE, George H. T. : *Tropical Africa,* Vols. I and II. 20th Century Fund, New York. $15.00, pp. 1073.

LAKDAWALA, D. T. & SANDESARA, J. C. : *Small Industry in a Big City—A Survey in Bombay.* (Publishers) University of Bombay. Rs. 20.00, pp. 384.

QURESHI, Anwar Iqbal : *Developments in Pakistan Economy Since the Revolution.* Nabeel Publishing House, Karachi. Rs. 5.00, pp. 195.

SHAH, C. H. : *Conditions of Economic Progress of Farmers.* Indian Society of Agricultural Economics, Bombay.

Futures Trading Seminar—History and Development. Vol. I. Mimir Publishers, Wisconsin. pp. 271.

Resources for the Third Five-Year Plan. Rs. 7.00, pp. 144. Indian Merchants' Chamber, Economic Research & Training Foundation, Bombay.

Inflation in a Developing Economy. Rs. 6.00, pp. 116. Indian Merchants' Chamber, Economic Research & Training Foundation, Bombay.

La Restauration des Monnaies Europeennes. Revue d'Economie Politique. (Publishers) Sirey. pp. 202.

The Soviet Seven-Year Plan. With introduction by Alec Nove. Phoenix. 10/6d., pp. 126.

The Soviet Union—Facts and Figures. Soviet News. 7/6d., pp. 131.

The Causes of the Great Depression, 1873-96*

The aim of this article is to consider the basic causes of the episode in economic history which has been termed the "Great Depression, 1873-96." A recent article by A. E. Musson [24] has summarised with great skill the mass of work which has been done on British experience during this period, but he did not attempt to present any definite conclusions on the basic international origins of the Depression. The problem is admittedly difficult because of the sketchy nature of most of the required statistical material, but it is believed that some progress towards a solution can be made even with present resources. What follows is therefore an attempt to sketch out the main outlines of a provisional solution.

I.

It is necessary to begin by defining the main problem for discussion. According to Mr. Musson "the most notable feature of the Great Depression was . . . the fall in prices, which so obsessed contemporaries" [24, p. 202], and the line to be followed here is that this secular decline in prices is the essential problem of the Great Depression. Since the severe price decline and the associated fall in interest rates was a world-wide phenomenon and not something peculiar to Britain [1] it would presumably be beyond dispute that the explanation of the price decline must be sought in trends in the international economy. It is possible that developments internal to the British economy were sufficiently important as to determine the world trends through the links of foreign trade and investment, but the evidence seems more consistent with a causal sequence in the reverse direction.

*I am indebted to Professor E. H. Phelps Brown, A. E. Musson and C. H. Feinstein for valuable comments in connection with this paper. Needless to say, they are not responsible for the views expressed here. Special thanks are due to Mr. Feinstein for making available a copy of his unpublished Ph.D. thesis.

[1]Data for five countries including U.K. is given in Phelps Brown and Hopkins [27].

The term depression is usually associated with severe declines in production levels and the emergence of abnormally heavy unemployment. For the Great Depression there is clearly no absolute decline in production to be explained, though in the case of the U.K. it has been shown that the average level of unemployment during the period 1874–95 was distinctly higher than that of the preceding and succeeding secular periods [24, pp. 201–2]. And it is noteworthy that the U.K. experienced the worst part of the Great Depression period in the 1880's, with an average level of unemployment for 1884–89 of 7·0%, the average for 1884–87 being 8·8%, a combination of severity and duration unequalled in any other depression during the period 1854–1913. This is something which requires explanation.

Can this particular problem be generalised for the other major economies? Direct data on unemployment is not available for the U.S.A. and Germany but indirect evidence suggests a negative answer. Professor Lewis and P. J. O'Leary [21] have presented charts of manufacturing production and related series for these countries which show levels for the 1880's which are well above the secular trend lines fitted to the period 1870–1913. They suggest that the 1880's were, for these countries, periods of secular boom associated with the boom phases of long cycles in residential building. For Britain, in contrast, the 1880's mark the depression phase of a long cycle in building. It is possible that levels of unemployment in the late 70's and early or middle 90's were relatively higher in the U.S.A. and Germany than in the U.K., but this can only be a conjecture based on differences in the phasing of the building cycle in the three countries.

Even in the case of the U.K. the apparent increase in unemployment levels in the Great Depression period almost disappears, given the rough nature of the indicators, if the period is extended to 1899 or 1900 [24, p. 201]. The same extension leaves the major part of the price decline intact though, of course, it does reduce the average annual rate of decline. All this suggests that the international secular price decline overshadows the local unemployment in the U.K. and differences in the pattern of building cycles, and stands out

as the major problem for analysis. It will be shown later that, given the explanation of the price decline to be proposed, the heavier unemployment in Britain and the complaints of falling profits can easily be fitted into place.

It is obvious from first principles that secular fluctuations in price levels must involve the alternation of secular periods during which there is expansion or contraction of effective demand relative to actual or potential supply of goods and services, demand and supply being interpreted as rates of growth for given price levels. The interesting question is that of deciding in particular cases whether the causal sequence originates from the side of demand or supply, or from some specific combination of the two, and it is here that disputes have arisen. In general, to judge from Mr. Musson's review, students have favoured explanations of the price decline of the Great Depression period which emphasise supply factors, e.g., the direct effects of falling transport costs and the outward shifting of supply schedules made possible by the development of primary production, in areas opened up as a result of transport innovations. On the demand side the popular theory has been the monetary theory, which argues from trends in gold production to trends in effective demand, either directly, or indirectly through the effects of gold supplies on banking systems and the creation of credit. The thesis of this paper is that the price trends of the Great Depression period can be understood only in terms of joint operation of the factors of demand and supply ; that greater importance must be given to the demand side in the origin of this secular price decline and that supply and demand factors were of more equal importance in reversing the secular price decline. The argument to be presented draws heavily on the pioneer work in this field by Professor W. A. Lewis [20] and Professor Phelps Brown and S. A. Ozga [28] and seeks to integrate their theories into an eclectic account of the causes of basic trends in the Great Depression period. It is not intended to present here a rigorous discussion of all the various theories of the price decline, but comment on some of them is necessary in connection with the argument that follows.

II.

The shortcomings of the monetary theory as an explanation of secular price movements in general have been ably discussed by Phelps Brown and Ozga [28] and it is unnecessary to repeat the details of their analysis. Of the 6 reasons which they advanced against the monetary theory 3 seem to be of over-riding importance; the lack of "systematic connection between changes in monetary stocks of gold and in the quantity of money in the hands of the public," the fact that "the price movements were not those of a *numéraire* but were regularly bound up with movements in the terms of trade" and the fact that "the long-term rate of interest moved not in the contrary direction to the price level but in the same" [28, p. 8]. The first of these points is derived from the well known analysis of J. T. Phinney [29] who found that, for the main countries between 1875 and 1913, "so many of the statistics examined show a relatively constant rate of growth of bank reserves and of bank currency, quite unaffected by variations in gold production or by trends in prices" [29, p. 677].

The validity of Phinney's findings have been questioned recently by R. P. Higgonet [13] who claims that the results may have arisen from using series in which inadequate coverage in earlier years raises the apparent rate of growth and so produces a misleading impression of constancy of growth [13, pp. 331, 353–4]. In relation to the U.K. Higgonet argues that "corrected data reveal a lower rate of increase of bank deposits during the period of falling prices (1874–95) than during the period of rising prices (1896–1914)" [13, p. 331].

It must be noted that, although Higgonet justifies this statement on the basis of an isolated estimate of bank money for 1873 [13, p. 347], the continuous series which he presents for the years 1877 to 1910 (linking with Tinbergen's series to 1913) does not support his thesis. From 1877–81 to 1893–97 the average annual growth rate of his series for total bank money, [13, p. 330, Col. 1], is 2·2%, which compares with a rate of 2·0% for the period 1893–97 to 1909–13, values which, given the nature of the data, could be used to support Phinney's constant growth rate. The average annual growth

rate of Net national income, using A. R. Prest's estimates [30], for the corresponding periods are 2·0% and 2·7% respectively.[1] The national income data also shows a smaller increase from 1873 to 1895 than from 1895 to 1913 as would be expected, the increases being of the same order of magnitude as those given by Higgonet for bank money. Comparisons based on data for single years are probably unwise, but those based on 4 to 5 year averages may, perhaps, be used tentatively. The average annual growth rates of bank money and national income given above seem to support the Phelps Brown—Ozga thesis that pressure of effective demand on the money supply was relaxing during the period of falling interest rates and was tightening during the period of rising interest rates [28, p. 8].

The behaviour of interest rates is the crucial argument against the monetary theory of the price decline. Whether or not the growth of bank money was constant from 1875 to 1913 it is not possible to argue the case for monetary stringency in the face of a persistent decline in the price of money. Nor can this fact be circumvented by shifting the argument into Wicksellian terms with the money rate falling, but falling less rapidly than the "natural" rate of interest. Such a move raises the question of the causes of the decline in the natural rate and shifts us into what is essentially a "real" analysis of the problem.

The role of transport costs will be considered next. It is undoubtedly true that a part of the general price decline of the Great Depression period was accounted for by improvements in the efficiency of transport but it is easy to exaggerate the importance of this factor. Examples of dramatic declines in transport costs can be given. From 1870–74 to 1893–97 the Caincross index of inward freight rates [4] shows a fall of 71%, whilst the cost of shipping wheat by steamer from New York to Liverpool fell by 69% [2, p. 130]. In the U.S.A. the freight

[1]For bank deposits the mean of Higgonet's figures for each year was used. For the years 1909–13 Tinbergen's estimates, as given by Higgonet, have been used since these estimates fall within the limits of Higgonet's figures from 1905 onwards. Since the deposit data are year-end figures the growth rates of income are calculated from the 4 year average periods 1878–81, 1894–97 and 1910–13. Revised estimates of Net national income by C. H. Feinstein [9] give growth rates of 2·0% and 2·3% for the two trend periods.

rate on wheat from Chicago to New York fell by 64% for lake and rail transport and 54% for rail transport [2, p. 130]. Despite such impressive declines there are grounds for limiting the importance of this factor. The following points have to be considered.

(1) The importance of transport costs cannot be assessed by looking simply at the record of falling money costs since part of the decline reflects the general fall in prices which is the problem to be explained. Certainly there was a genuine fall in real costs of transport during the period 1873–95, but this was no sudden development. The fall in real costs (and even money costs in some cases) continued well after the 1890's, whilst real costs must have been declining from 1850. In view of the well-known volatility of freight rates, the apparent fall in real costs over the Great Depression period may well have been exaggerated by an excessive decline in money costs, which can be explained by our later argument in terms of changes in the growth rate of effective demand.

(2) Leaving aside the argument of (1) there are still substantial declines in prices, net of money costs of transport, which have to be explained. Wheat prices are an interesting case. It is easy to show that for this commodity a major part of the decline in U.K. import prices was accounted for by falling money costs of moving wheat [7 and 15]. Yet the U.S.A. farm prices fell by about the same percentage as the U.K. import prices from 1870–4 to 1893–7, the declines being 46% and 45% respectively. [Using data in 15, p. 656]. Even in the case of the States representing the cheapest centre of production, Hooker's "6 wheat States" and Nebraska, the declines were 28% and 30% over the same period [15, p. 656]. Moreover examination of the relevant freight cost data shows that the major fall in freight costs had occurred by 1881. From 1870–4 to 1881–5 the average cost of moving wheat by lake and rail from Chicago to New York fell by 3s. 7d. per quarter and the cost of steam freight New York to Liverpool by 3s. 0d. In the same period the average cost of wheat imported from the U.S.A. into the U.K. fell by 10s. 3d. From 1881–5 to 1893–7 the same freight costs fell by a total of 1s. 9d. whilst wheat fell by a further 14s. 7d. [Data from 2, p. 130].

More generally, indexes of export unit values (f.o.b.) by C. P. Kindleberger [17, p. 157], for some important exporters of primary products, show falls from 1873 to 1895 of the same order of magnitude as the fall in the U.K. import unit value index. The Australian index fell 48% and the Indian and New Zealand indexes fell 32% compared with 39% for U.K. imports. Between 1879 and 1895 U.S.A. export values fell 29% against 26% for the U.K. The Canadian index provides an exception showing a fall of only 2½% from 1873 to 1895.

(3) If the decline in transport costs was the only, or major, factor causing the secular price decline it is difficult to explain the different trends in food and raw material prices in the 1870's and early 1880's. By 1882 raw material and food import prices given by Schlote [33, pp. 176–7] had declined by 21% and 9% respectively from their levels in 1873. The incidence of transport costs in these sectors cannot have been so different as these figures would suggest. The behaviour of the two sets of prices indicates that other determinants of the demand and supply schedules were more important than the transport cost element.

The points listed suggest the need for a cautious appraisal of the role of transport costs in the secular price decline. For relatively cheap and bulky commodities, like wheat, they may well have been a major factor but it is not clear that the argument can be generalised for all commodities.[1] Also, the dramatic decline in money costs of transport cannot be regarded as exclusively the result of improvements in the efficiency of transport.

A more general supply side explanation of the price decline of the Great Depression is that which postulates

[1]An extremely rough indication of the relative importance of transport costs can be got from estimates of world imports (c.i.f.) and world exports (f.o.b.) for the periods 1872–3 and 1895–99 as presented by Imlah [12, p. 189]. For 1872–3 the c.i.f.–f.o.b. difference was 11·1% of the value of total imports; for 1895–99 the difference was 7·9% of the value of imports. Using Imlah's price index for U.K. imports as representing the world index the volume of world imports for 1895–99 at 1872–3 c.i.f. rates would have cost 10·4% more whilst the volume of world imports in 1895–99 at 1872–3 prices would have cost 65% more. From this we may deduce that money transport costs can only explain about one-sixth of the overall decline in the price of imports.

excessive growth of supplies of primary products and even manufactures arising from the development of new areas of production. To quote from Layton and Crowther [18, pp. 87–8]: "it appears that during the 'seventies, 'eighties, and early 'nineties the expansion in the world's output of all kinds of commodities far outran that of any previous period in history. While the New World began to flood European markets with food products and raw materials, the arts of production made enormous strides ; manufactures were, moreover, stimulated in both Germany and the United States by the opening up in both countries of new sources of iron and coal . . . " A similar argument with the emphasis on primary products has been put forward by Rostow [32] who writes of the trend periods of falling prices as being caused to a large extent by "prolonged overshooting, in particular sectors of the economy, and by the reaction from such excessive distortions in the sectoral pattern of growth . . . the overshooting process . . . arose from eras of war and/or of large-scale expansion in capacity for raw materials and foodstuffs, usually connected with the opening up of new areas" [32, pp. 126, 129].

This emphasis on the supply side has been challenged by Phelps Brown and Ozga in a general analysis of the causes of secular price movements [28]. They present graphs of total output of some primary products for varying periods between 1830 and 1950 and argue that these fail to show any significant variations in their growth trends such as would account for the turning points in the price trends [28, p. 8]. They conclude that "the inference is strong that the general movements of prices arose from changes in the rate of growth not of supply but of demand" [28, p. 10]. They propose, therefore, a theory of secular price movements in terms of the following mechanism. The world economy has been subject to periodic changes in the growth rate of effective demand arising from variations in the growth rate of industrial capacity. "When industrial capacity tends to grow more quickly than the output of primary products, a shortage of these products is felt . . . the industrial economies experience a rise in wholesale prices. This spreads through the economy : rises in the cost of raw materials raise the prices fixed for final products ; the rise in the cost of living moves

wage-earners to press for rises in money wages ; with more or less lag, other money incomes are raised to maintain customary differentials, and this causes further rises in costs and prices" [28, pp. 13–4]. They emphasise the effect of changes in the wholesale price level of basic commodities on the formation of expectations regarding future prices which in turn affects the price policy of sellers [28, p. 14]. When industrial capacity grows more slowly than output of primary products the arguments are reversed to produce a fall in the general price level.

On the question of repercussion effects of excess demand and supply in the markets for primary products on the general price level these alternative theories probably stand on common ground. They differ in their specification of the basic causes of the alternating conditions of excess demand and supply and here the argument must be settled by reference to historical fact. Their respective merits in relation to the Great Depression period will be examined from this standpoint in the following section.

III.

In the case of industrial or manufacturing production the Layton-Crowther theory seems to be directly contradicted by such evidence as is available on trends in world production. Instead of an acceleration of production after the early 70's there is a strong suggestion of deceleration. The demonstration can begin by considering the position of the U.K. Between 1850–4 and 1870–4 the average annual growth rate of industrial production (excluding building) as measured by Hoffmann's index [14], was about 3·0% compared with 1·5% for the period 1870–4 to 1893–7. Since the U.K. accounted for almost one-third of world manufacturing production in 1870 and about one-half of world output of coal [19, p. 13 and 16, p. 365] this marked change in the U.K. growth rate would lower the growth rate of world production by some 0·5% per annum in the absence of any corresponding acceleration in other countries.[1]

[1]To minimise computing work the many trend rates of growth used in this article have been calculated, where possible, from appropriate 5 or 10 year averages of annual values. From inspection of graphs and alternative test calculations the writer is satisfied that the growth rates would not differ materially from those used if they had been derived from linear trend lines fitted to the logarithms of annual values. Cotton consumption provides a possible exception.

For the U.S.A. and Germany, which together accounted for a further 35% of world manufacturing in 1870, indexes of manufacturing or industrial production are available only from 1860. In neither case is there any evidence of acceleration over the period 1873 to 1895 ; rather both series are consistent with deceleration, less obvious in the case of Germany, though cyclical phasing of the period 1860 to 1874 exaggerates the trend in the first period. Frickey's index of manufacturing production [10] shows average annual growth rates of 5·5% between 1860–4 and 1870–4 and 4·3% between 1870–4 and 1893–7. For industrial and commercial production the deceleration is greater, the respective growth rates being 6·5% and 4·7%. Wagenführ's index is not easily available but is published in graph form in Schumpeter [34, p. 467].[1]

Over a longer period there is evidence of deceleration in strategic series. For the U.S.A. total coal production grew at an average annual rate of 8·4% from 1850–4 to 1870–4 compared with 6·0% from 1870–4 to 1893–7 [Data from 40, p. 142]. Schumpeter [34, pp. 488, 486] gives charts of railway freight ton-miles and pig-iron consumption from the early 50's which show definite declines in trend growth rates after 1873–4. For Germany pig-iron consumption shows the same behaviour [34, p. 486], whilst Wagemann's estimates of pig-iron consumption *per capita* and soft coal consumption *per capita* show drastic decelerations in growth, the average annual rates being 7·5% in each case from 1851–5 to 1870–4, falling to 3·4% for iron and 3·0% for coal from 1870–4 to 1893–7 [Data from 41, p. 265].

For the three major industrial nations which together must have represented the major part of world production the evidence suggests deceleration of industrial growth during the period of the Great Depression. This suggestion is confirmed by several strategic series which have substantially world-wide coverage. In 1934 Carl Snyder [36] produced charts showing estimates of world railway freight tonnage, industrial employment (excluding agriculture) for 13 countries, and computed total trade and production, each of which shows a definite

[1]Wagenführ's index has been criticised by Lewis [21, p. 149], as showing insufficient growth between the 1870's and 1880's. Even allowing for Lewis's proposed correction it is doubtful if the index would show acceleration of trend after the early 70's.

break in the trend rate of growth around 1870. World output of coal and pig-iron show the same deceleration. World output of coal grew at an average annual rate of 5·2% from 1851–60 to 1869–73 falling to 3·7% from 1869–73 to 1894–8. World output of pig-iron grew at 5·3% between 1851–60 and 1870–4 falling to 3·3% from 1870–4 to 1893–7.[1] In the case of pig-iron the trend from 1870 to 1895 should be raised to allow for the economies effected by the conversion of pig-iron into steel. If the total for 1893–97 is raised by 20% of world output of steel for that period [from 3, p. 275] the adjusted growth rate becomes 3·8%. Thus world coal and pig-iron production agree to a deceleration of growth of the order of 1·5% per annum after 1870–4. Deceleration of growth can also be seen in world output of wool, jute, lead and zinc.[2] Copper production accelerates after 1870–4 and cotton consumption in Europe, U.K. and U.S.A. though greatly disturbed by the Cotton famine period shows a decline in growth from 3·5% for the period 1846–50 to 1871–5 to 3·0% for the period 1871–5 to 1896–1900.[3]

The evidence presented here may be interpreted as supporting the concept of a Kondratieff long wave in economic growth between 1850 and 1895. There would seem to be a good *prima facie* case for the view that the trend growth rate of world industrial production fell by a minimum of about 0·5% per annum between the early 1870's and middle 1890's. If the

[1]Estimates for coal and iron for 1851–60 from Mulhall [23, pp. 119,334]. Coal 1869–73 and 1894–8 from Jevons [16, p. 365]. Iron for 1870–4 and 1893–7 from Burnham and Hoskins [3, p. 273].

[2]Based on the graphs presented by Phelps Brown and Ozga [28] and data from Mulhall [23].

[3]Based on [23] for copper production and [2, p. 443] for cotton consumption. The deceleration of growth of metal and coal production is confirmed by component series of Snyder's estimates of world output of basic commodities valued at 1923–5 average U.S. prices [33, p. 13]. The sub-series "metals and mineral products" shows growth rates of 5·1% for 1850–4 to 1870–4 and 4·2% from 1870–4 to 1893–7. Snyder's sub-series for textiles and rubber shows an acceleration of growth after 1870 of 0·8% per annum. Between 1850 and 1883 this series covers only cotton, wool and rubber (the last item must have been negligible). The behaviour of this component is contradicted by the other evidence relating to cotton and wool, discussed above, and does not seem trustworthy. I am indebted to the Union Corporation Ltd. for supplying unpublished data relating to the Snyder series.

size of the deceleration in coal production and adjusted pig-iron production is taken as a plausible upper limit to the size of the overall decline we can say that the trend rate of growth of world production fell by some 0·5%–1·5% during the period of the Great Depression.

In the face of a decline in the growth of industrial demand of this order even an unchanged rate of growth of supply of primary products could be expected to give rise to a general decline in prices via the process described by Professor Phelps Brown and Mr. Ozga. It is necessary now to consider whether the process was exaggerated by an acceleration of growth of supply as visualised in the Layton-Crowther-Rostow theories.

In considering trends in the supply of primary products a distinction must be made between raw materials production and food production. The behaviour of raw materials production has, in effect, been considered already, partly because of the close technical relationship between industrial production and raw material inputs, and partly because indexes of industrial production are, in practice, constructed substantially on the basis of consumption of raw materials. But a deceleration of growth of actual production of raw materials need not imply a deceleration of growth of *potential* output of raw materials, since the actual output adjusts rapidly to the demand arising from industrial production.[1] It is hardly possible to measure directly the growth trends of potential raw materials production for the period under discussion but some insight can be gained by indirect methods, using a technique developed by Professor Lewis [20]. Professor Lewis was able to deduce the course of potential supply of raw materials by using data relating to the terms of international trade between raw materials and manufactures and an index of world manufacturing production. He found that the assumption of constant trend growth rates of potential raw materials supply for the periods 1880–1913 and 1921–38 enabled him to predict the course of the raw materials terms of trade. For the inter-war period the method gave a good prediction of the terms of trade but for the period 1870–1913 the equation with parameters fitted for the years 1880–1913 gives a serious understatement of the value of the terms

[1]See Lewis [20, p. 119].

of trade between 1870 and 1885 [20, p. 121]. This error can be eliminated by recalculating the estimating equation for the full period but the error is then shifted to the 1890's whilst the pattern of residuals indicates clearly the existence of a U-shaped trend factor.

If Professor Lewis's method is modified to allow for a non-linear growth of potential supply of raw materials the resulting equation enables a better prediction to be made of the raw materials terms of trade between 1870 and 1913. The mean error of estimation is reduced and the trend factor in the residuals is virtually eliminated.[1] The equation carries the implication that potential supply of raw materials was increasing at a slightly higher average rate between 1870 and 1895 than between 1895 and 1913, whereas actual supply must have varied in the opposite way.

It is interesting to use the equation to test various assumptions about the growth of world manufacturing production and potential raw materials supply for the period before 1870. Data for calculating the raw materials terms of trade on the Lewis basis [20, pp. 117, 137–8] do not exist for years before 1854 and the behaviour of cotton prices during the early 60's distorts the terms of trade for those years. The period 1854–60 is therefore selected for the tests. The average value of the terms of trade for this period was 100·5 (1913 = 100) compared with 101·3 for the period 1870–4. It has been suggested that world industrial production may have grown

[1]The equation is as follows : log P_{RM} = ·78450 + ·70009 M_Q — ·01501 N + ·000085 N^2, where P_{RM} is Lewis's index of the raw materials terms of trade, M_Q the index of world manufacturing production (1913 = 10·0 in each case) and N is the number of years elapsed since 1871. Five year moving averages were used as in the original work since we are only interested in trend relations. The index of world production is taken from [19] and includes Russia. The coefficient of multiple correlation is 0·964, whilst the mean deviation of actual observations from the regression estimates is 0·76%. It should be pointed out that the equation given here, like those of Professor Lewis, depends for its validity on the assumption that the price elasticity of demand for raw materials is zero. The coefficient of M_Q is, therefore, the reciprocal of the short-run elasticity of supply whilst the coefficients of the N terms show how much price would be reduced each year given an unchanged level of manufacturing production (i.e. demand). This fall in price is determined by the annual increase in raw materials production, the annual fall in raw materials consumption per unit of output and the elasticity of supply [20, p. 120].

between 1850 and 1874 at a trend rate of 0·5%–1·5% greater than the rate for 1870–4 to 1893–7. Taking these limiting rates for manufacturing production also (i.e., 3·6% and 4·6%) the estimating equation gives values of the terms of trade for 1854–60 of 123·2 and 111·3. To eliminate the error in the estimate of the terms of trade would require a growth rate of world production of some 5·5% per annum before 1870, which is undoubtedly excessive. It is clear from these results that the non-linear trend of potential raw materials supply between 1870 and 1913 cannot be extrapolated backwards to 1854. The rate of growth between 1854 and 1870 must have been smaller than the rate of growth between 1870 and 1895.

Suppose we assume that between 1854–60 and 1870–4 the growth of potential supply of raw materials was equal to the average growth rate between 1871 and 1913.[1] Then for assumed growth rates of manufacturing production of 3·6% and 4·6% per annum the equation gives estimates of the terms of trade for 1854–60 of 105·5 and 95·4 respectively compared with an actual value of 100·5. For an assumed growth of manufacturing of 4·1%, which is probably near to the actual trend rate for the period the estimated terms of trade is 100·3, an error of prediction of only 0·2% of the true value. Since this range of errors seems very reasonable it seems in order to conclude tentatively that the rate of growth of potential supply of raw materials increased during the Great Depression period. It must be noted that this apparent increase in the growth rate of potential supply may well reflect the factor of transport costs whose influence would be merged into the coefficients c and d of the estimating equation. The combined effect of the changes in the supply side (growth of supply and falling transport costs) on the trend in the terms of trade must have been modest in size at least over the full period of the Great Depression. The pre-1870 growth of supply would have reduced the terms of trade by 2·7% per annum given an unchanged level of manufacturing production. For the full period 1870–95 the average reduction would be 3·0%. If manufacturing production

[1]This means a growth rate of supply which would cause a decline in the terms of trade between 1871 and 1913 equivalent to that caused by the assumed rate of growth of supply in the estimating equation.

decelerated by about 1% the demand factor would account for a change in the trend of the terms of trade of —0·7% per annum compared with —0·3% for the supply side.[1] If the two factors operated with similar orders of importance on the absolute price level the Phelps Brown–Ozga thesis is substantially, but not completely, vindicated.

In the case of food production the difference between actual and potential growth rates of supply is less important than with raw materials since there is reason to suppose that the rate of food supply adjusts to changes in demand with greater inertia and may even adjust in a perverse direction.[2] It is noteworthy that world food production continued to rise in the great depression of the 1930's in the face of the substantial decline in industrial production and incomes. The trend of actual food production can therefore be taken as roughly equivalent to the trend of potential supply.

The nearest approach to an index of food production for the period since 1850 is the food component of Snyder's index of "products of the soil" [35, p. 13]. The coverage of this index, which has been published only in graph form is meagre before 1870.[3] For what it is worth the index provides no support for the thesis of a major acceleration of growth of food production between 1870 and 1895, the average annual growth rates being 2·5% from 1850–4 to 1890–4 and 2·6% from 1870–4 to 1893–7, though it does suggest that food production was high relative to the 1850–1900 trend line in the 1880's and low in the early 70's and 90's.

The general behaviour of the Snyder index is consistent with the record of growth of some important food products over the period 1850–1897. Growth rates for the two trend periods are given in the table below.

[1]The actual change was —0·84% per annum the difference being errors of estimation and rounding.

[2]See Lewis [20] p. 122 and Ojala [25] p. 179.

[3]Between 1850 and 1870 it is limited to wheat, corn and sugar. Between 1870 and 1895 only coffee is added. Data and details of coverage were kindly supplied by the Union Corporation Ltd.

AVERAGE ANNUAL GROWTH RATES OF WORLD OUTPUT OF SELECTED FOOD COMMODITIES

Commodity	Trend Period	Growth Rate (%)
1. Wheat	1851/60 – 1871/80	2·1
	1871/80 – 1894/99	2·1
2. Other Grains ...	1851/60 – 1871/80	1·4
	1871/80 – 1895/99	1·9
3. Meat (Beef, Mutton and Pork) ...	1851/60 – 1874/84	1·0
	1874/84 – 1891/96	1·0
4. Sugar	1850/54 – 1870/74	3·8
	1870/74 – 1893/97	4·2

Sources : (1) Mulhall [23] p. 8 for 1851/60 and 1871/80, Malenbaum [22] p. 238 for 1894/99. Mulhall gives 2 estimates for 1871/80 without indicating which is to be preferred. If the higher estimate had been used the growth rates would have been 2·4% and 1·7%.
(2) Mulhall [23] p. 283 and [39]. For second trend period the figure covers total of rye, maize, oats and barley ; for the first period "sundries" is added.
(3) Mulhall [23] p. 283 and p. 714. The total for 1891/96 is not a genuine average but appears to be an average of estimates for the different countries for varying periods between 1891 and '96. For the calculation of the growth rate 1893 was taken as the mean year.
(4) Deerr [8] pp. 490-1. Deerr's figures for total output do not appear to make the best use of his data for individual countries but corrected totals would show only trifling changes for the periods used here.

The commodities listed above, though few in number, may have accounted for as much as 40% of total food and drink expenditures exclusive of retail margins and taxes.[1] Doubtless the earlier estimates are subject to a substantial margin of error but in the absence of more reliable estimates they merit attention. It will be noted that there is no evidence of acceleration of growth in the important cases of wheat and meat production. The case of sugar production is interesting since

[1]Based on Mulhall [23] p. 286. The estimate relates to 1887 and covers expenditure in Europe, U.S.A., Canada and Australia.

this commodity was noted by Layton and Crowther [18, p. 89] as the one undergoing the greatest price decline between the 70's and 90's. The increase in the growth rate of sugar production is explained largely by the expansion of subsidised beet production in Europe rather than the expansion of production following the transport revolution.[1]

It would seem, then, that with the statistical evidence at present available we can rule out the theory that the Great Depression period witnessed a major acceleration of the growth of potential production of primary products. The relatively small acceleration which occurred together with the decline in real transport costs cannot explain a price decline of the required magnitude. The Phelps Brown–Ozga theory, with its emphasis on the trends of effective demand seems better fitted to explain the price decline, though it cannot provide a complete explanation. The continued growth of primary production between 1870 and 1895, at rates equal to or somewhat greater than those obtaining before 1870, makes Rostow's theory of prolonged overshooting formally valid, but primary production would not have been greatly excessive if industrial production had maintained its pre-1870 trend rate of growth.

Why did industrial production and effective demand decelerate in growth after 1870? In a previous article [5] I have suggested that the U.K. "climacteric" of the 1870's, the decleration of growth of production and productivity, was caused by a decline in the trend rate of growth of U.K. exports, which in turn was associated with the emergence of the U.S.A. and Germany as serious industrial competitors. But a switch of industrial advantage from the U.K. to the U.S.A. and Germany need not have involved any deceleration of growth of the three countries together.[2] Nevertheless there is evidence that the three economies combined did experience a deceleration of industrial growth. Some part of the decline was undoubtedly

[1]Deerr [8, pp. 490–1] gives figures showing cane as 60% of total in 1870 and 40% in 1895.

[2]It is interesting to notice that the combined share of the 3 countries in world manufacturing production and exports of manufactures shows little change over the period 1870 to 1913. Data in [19] give the joint share in manufacturing as 68·3% in 1870 and 65% in 1913. In exports the joint share was 60% in 1881/5 and 58% in 1913.

the endogenous result of the deceleration itself; the product of decelerating industrial demand for primary products in world trade and the worsening terms of trade for primary producers interacting through the mechanism explained by Professor Lewis [20] to produce a relative stagnation of demand for exports of manufactures during the Great Depression period.[1]

For the independent basis of the deceleration of industrial growth and demand three factors may be suggested. The first, and probably most important, was the change in the trend of railroad investment after the 1870's, especially in Europe and the U.S.A. By 1873 the main outlines of the railway system had been developed further expansion being of a secondary character. From 1850 to 1870 world railroad mileage increased at an average annual rate of 8·8%, from 1870 to 1897 the rate was 4·7%.[2] A decline in the growth rate of such a highly capital intensive industry provides a strong basis for a decline in the growth rate of real investment and effective

[1]Lewis's argument has been criticised by Musson [24, pp. 219–20] on the grounds that it assumes that international trade consists of the exchange of manufactured goods against primary products whereas the greater part of trade in fact consists of the interchange of primary products and of manufactures. This is a misunderstanding of Lewis's argument which is based on the existence of a constant share of manufactures in world trade as an observable but unexplained fact. The key to this problem may lie in the fact that the industrial nations export primary products as well as manufactures and substantially determine the level of world demand for both. Thus an expansion of industrial demand in, say, the U.S.A. would lead to an increase of U.S. imports both of manufactures and of primary products and the expansion of world demand induced by this would lead to higher demand for manufactures and primary products including those exported from the U.S.A. Thus world trade would expand simultaneously in all three sectors without any necessarily great change in the share of manufactures in the total but with a net change in the terms of trade between manufactures and primary products. All this is not to deny that tariffs and industrialisation may have played some part in checking the growth of exports of manufactures as Musson suggests, but this argument has to confront the accelerated growth of trade in manufactures after the 1890's.

[2]Based on Mulhall [23] pp. 495, 795. The decline was equally valid in terms of mileage added per decade.

demand.[1] Secondly, the application of iron and steam to shipbuilding was a capital saving innovation. The cost of steamships per ton was higher than that of sailing vessels but the carrying power ratio was even greater. This factor would operate most effectively after 1870 when the world tonnage of sailing vessels ceased to grow. The volume of world steam tonnage also shows deceleration after 1870 with average annual growth rates of 7·3% from 1850–70 and 5·8% from 1870–96.[2] The result of this deceleration can be seen in the trend of U.K. steam shipbuilding which must have accounted for the major part of world output at this time. Growth at an average annual rate of 13·7% from 1858–62 to 1870–74 was followed by an average rate of 2·0% from 1870–74 to 1893–97 [Data from 2, p. 379]. Thirdly, there seems to be a case for allowing the familiar argument that the concentration of wars in the period 1850–1873 was important in determining the trend of effective demand and the price history of the period, relative to what followed.

The three factors mentioned are visualised as operating to produce a period of secular expansion of investment and effective demand which was not maintained in the Great Depression period. Statistical support for a decline in investment growth during the Great Depression can at present be given only in the case of the U.K. and the U.S.A. Feinstein's estimates of U.K. domestic real investment since 1856 show a

[1]Ulmer [38, p. 11] gives U.S. railroad gross investment as 20% of total gross investment in the 1870's. Feinstein [9, pp. 91, 103] gives U.K. railways 22% of gross domestic fixed investment for the period 1856–69. Feinstein uses this factor *inter alia* to explain the trend of U.K. real investment after 1870. It seems possible that the trend in railroad investment in the U.S.A. was matched also by a similar trend in capital formation in residential building. The two industries show strong related long cycles—the "transport-building cycle." Riggelman's index of real building permits *per capita* from 1831 onwards (as given in [37], p. 298) does not show any deceleration of trend after 1870, but the limited coverage of the index in the earlier years is likely to distort both trend and amplitude of fluctuations. In view of the importance of the demographic factor in housing investment and the decelerations in U.S.A. population growth after 1860 and 1880 the proposition advanced here seems reasonable.

[2]Data from Mulhall [23] pp. 522, 801. Tonnages for U.S.A. and Spain have been converted roughly from a gross to a net basis for the calculations, using a ratio of 1·6 : 1.

pronounced change in trend after 1877.[1] For the U.S.A. Goldsmith's data on the growth of total reproducible tangible wealth in 1929 prices, excluding consumer durables [11, p. 307], show average annual growth rates of 5·4% from 1850 to 1880 and 4·8% from 1880 to 1900. In view of the difference in the trend behaviour of industrial production after 1870 in the U.K., the U.S.A. and Germany, it seems that the main inpact of the Kondratieff downswing in effective demand was borne by the U.K. and probably Germany was least affected. Britain's loss of competitive power to the U.S.A. and Germany after 1870 therefore merely exaggerated a change in economic growth which was in any case inevitable. The impact of the Kondriatieff downswing was felt by U.K. exports when the Juglar boom collapsed in 1872–3, but home investment shows a later turning point because of the influence of investment in residential building. The building boom though controlled substantially by demographic factors may have been prolonged by the release of finance following the collapse of foreign investment [9, Chap. 6].

It is worth noting that the hypothesis presented here, if valid, does not imply that the growth of effective demand was excessive in the period before the 1870's. Although prices rose sharply from 1852 to '54 the trend level from then until 1871 was virtually horizontal in the U.K. and Germany [18, p. 240].

The falling level of interest rates during the Great Depression can be accounted for by a combination of real and monetary factors. The assumed change in trend of real investment relative to the overall propensity to save would initiate the downward trend and this would be reinforced by the deceleration of effective demand in the face of the tendency for money supplies to grow at a relatively constant rate, or at least at rates in excess of the needed growth for transactions purposes.

The different behaviour of raw material and food prices in the 1870's which was noted earlier in the discussion of transport costs can be accounted for by differences in phasing

[1]See [9] Fig. 2.4 p. 102. It must be emphasised that Feinstein's estimates for the earlier years do not claim great precision.

of the demand and supply changes of the Great Depression period. If the general shape of the food component of Snyder's index of "products of the soil" is accurate, food production declined temporarily below trend levels in the early 1870's and then grew more rapidly than the trend rate until the middle 80's. At the same time demand for food would be sustained until the later 70's by the important U.K. component which was influenced by the pattern of the residential building cycle. The volume of food imports to the U.K. was rising both absolutely and relative to trend until 1879 [21, p. 123 and 42, p. 327]. In contrast U.K. imports of raw materials were declining relative to trend during the 1870's showing a close relationship with industrial production excluding building [21, p. 121]. The overall demand for raw materials is likely to have collapsed fairly early in the 1870's with the coincidence of the Juglar slump in the main countries. This same factor may well explain a large part of the decline in money costs of transport in the 1870's in conjunction with supply lags in shipbuilding and would, of course, have secondary repercussions on the price-level through the effect of falling transport costs on raw materials prices.

IV.

To complete the analysis of the secular price decline it is necessary to discuss briefly the nature of the forces which produced the transition from deflation to inflation in the world economy after 1895. The Phelps Brown–Ozga theory requires an increase in the rate of growth of industrial demand relative to the growth of supply of primary products. This would turn the terms of trade against industrial products and cause an increase in the absolute level of primary product prices, which would then have direct and indirect repercussions on the prices of industrial products and goods in general. Such a process had operated, albeit briefly, during the Juglar booms ending in 1882 and 1889 ; consequently the actual turning point of prices and the terms of trade in the mid-90's does not present any special problem. What has to be explained is the failure of prices to resume their downward slide when the Juglar boom of the 90's was completed.

A proximate explanation of major force is to be found in the accelerated trend growth of world production from the mid-90's until 1913. According to the League of Nations index [19] world production, which had grown at an average annual rate of 3·1% from 1870–74 to 1893–97, attained an average growth rate of 4·3% per annum from 1893–97 to 1909–13, a significant change in trend.[1] It seems, however that the demand factor cannot provide a complete explanation of the acceleration of prices. In the case of raw materials this judgment is based on the relative influence of demand and supply on the changed trend in the terms of trade. This amounted to an average annual net increase of 1·5% of which some 0·9% can be attributed to the demand factor on the basis of the estimating equation. The net change attributable to growth of supply is 0·7% the difference of 0·1% being error of prediction. In the case of food production the demand side is at first sight more important than the supply factor since the food component of Snyder's index of the "products of the soil" shows virtually identical trend values over the two sub-periods (2·6% and 2·5%). It is noticeable, however, that its growth rate falls off after 1900 (2·2% per annum 1900–04 to 1909–1913), and in these years the supply factor must have become increasingly important. The causal linkage between primary product prices and the general price level is likely to be more immediate and definite for raw materials than for food and in view of the difficulty of making precise statements, given the quality of the basic statistical data, it seems prudent to allow demand and supply roughly equal influence in reversing the fall in prices.

It is not altogether clear why world production should have increased its trend rate of growth after the 1890's. Some part of the increase was undoubtedly endogenous. Given the turning point in production and the terms of trade, what may be termed the "Lewis effect" would raise the growth rate of

[1]Note that the index is based on weights for the period 1925–29 which gives the U.S.A. 42·5% and only 20% to the U.K. and Germany combined. Experiments with different weighting systems suggest that the trend growth rate is slightly exaggerated by the heavy weight given to the U.S.A. but that the difference in trend rates for the two sub-periods is still substantial. Using 1896–1900 weights the acceleration in growth would be reduced by about one-fifth of the value given in the text.

primary producers' incomes and so world exports of manufactures and world production [20]. The change in the trend of world exports is strikingly illustrated in the League of Nations study [19, p. 17]. But this presumes the existence of more basic forces to produce the turning point.

Professor Phelps Brown and Mr. Ozga have put forward an ingenious argument to explain the turning point. They point out that "if two series growing at different constant rates are added together, their sum grows at an increasing rate" [28, p. 10], a theorem which could result in an accelerated trend of world production even if rates of growth were constant in the different countries. The shift in the shares of world production of the U.K., U.S.A. and Germany between 1870 and 1913, which is associated with their different trend growth rates [1] suggests that the theorem may be applied to the present problem. It is doubtful, however, if the mechanism can provide a complete explanation of the basic change in trend in world production; its influence seems to be of the order of 0·3% per annum.[2] In fact, it seems clear that individual countries did not have constant trend growth rates of production between 1870 and 1913. The U.S.A. shows growth rates of 4·3% per annum from 1870–74 to 1893–97 and 5·6% for the second period. Other countries show the acceleration to a greater or lesser degree. For countries heavily dependent on exports this may well follow from the operation of the "Lewis effect" reinforced by the deceleration of growth of primary products. But it seems difficult to avoid the conclusion that a substantial part of the acceleration in world production stemmed from the secular swing in U.S.A. production which was transmitted to the world economy via U.S.A. imports [21, p. 117 and 6, pp. 25–6]. The secular swing in U.S. production is largely accounted for by the phasing of the transport-building cycle [21] [6]. Given

[1] In 1870 the U.K. share was 32%, U.S.A. 23% and Germany 13%. Following growth rates of 1·6%, 5·0% and 4·7% respectively the shares in 1913 were 14%, 36% and 16% [19].

[2] Based on the following calculation. The countries comprising the world index [19] were divided into 6 groups with similar trend growth rates between the individual years 1870 and 1913. Output in 1870 for each group was based on the 1870 shares of world production [19, p. 13] and was then assumed to grow at constant rates equal to the 1870/1913 trend values. Group outputs were then summed for 1895 and 1913. The resulting trend growth rates for the aggregate were 3·9% (1870–1895) and 4·2% (1895–1913).

the joint operation of these factors we may theorise that there was an acceleration in the growth rate of world incomes, which after 1895 were rising more rapidly than the world supply of money, the consequent pressure on inactive balances explaining the change in the trend of interest rates.

V.

Two final problems requiring discussion relate to the special position of the U.K. during the Great Depression period, the abnormally high level of unemployment and the declining profits of the period.[1] The explanation of the higher level of unemployment is fairly simple ; it follows naturally from the deceleration of real effective demand in the face of a relatively constant growth of occupied population, the deceleration in demand being explained by the largely autonomous decline in the growth of exports, with its repercussions on the rate of investment. A further influence on the trend of investment was the long downswing in the cycle in residential building and investment by local authorities which were both substantially influenced, directly and indirectly by demographic changes. It is significant that unemployment was at its highest in the 1880's when residential building had fallen to its lowest levels. It is true that the lower level of home investment in the 1880's was compensated for by an increase in foreign investment, but the compensation was incomplete, probably because the boost to exports given by foreign investment could only mitigate the effects of the climacteric in export growth.[2] In fact, given

[1]For a detailed discussion of the question of profits see Feinstein [9].

[2]Rostow's earlier theory that the "central causal force in the Great Depression was the relative cessation of foreign lending" [31, p. 88] has been soundly criticised by Musson [24], Thomas [37, p. 188] and Feinstein [9, pp. 269–70], though his "less precise" reformulation "What Happened when the Railways were Built" [31, p. 88] is consistent with the argument of this article. Musson finally concedes a part of Rostow's case on the grounds that the percentage of net foreign investment to national income was lower during the period 1875–96 (2·7%) than during the period 1870–74 (5·6%) and 1897–1913 (4·1%), whilst the percentage of home investment was much more stable in the range of 7–7½% [24, p. 217]. This concession was probably unnecessary. The period 1870–74 is too short to be relevant and Feinstein's data from 1856 onwards indicate that the relative share of foreign investment declined less than that of fixed home investment during the Great Depression. The actual values are as follows (net fixed home investment first) : 1856–74 5·9%, 4·8% ; 1875–96 5·2%, 4·6% ; 1897–1913, 6·0%, 5·5%. (Percentages are based on sums for each period, not averages of annual values). Thus the Great Depression saw a decline in the percentage of total investment to net national income.

the trend in exports and investment, what is surprising is not that the average level of unemployment increased but that it increased so little.

The absence of more serious levels of unemployment can be accounted for by the change in the trend growth of industrial productivity after the 1870's. The average annual growth of labour productivity in the industries covered by the Hoffman index of production, which had been not far short of 2·0% from 1850 to 1870 had declined to a rate of 0·3% from 1875 to 1899 [5, p. 7]. If the growth of labour productivity had maintained its pre-1870 rate the labour force needed to sustain the level of production in the mid-90's would have been smaller than that required in 1871 whereas the actual force was some 30% greater! Of course, this calculation presupposes a continuance of the pre-1870 growth rate of capital per head and the annual increment due to changes in technique and organisation which seems to be such an important determinant of the rate of economic progress [1]. It is doubtful if the rate of capital accumulation per head in manufacturing industry fell sufficiently after 1870 to account for the decline in labour productivity growth by itself. Consequently there is a strong hint that labour was used less efficiently or that the gains from changes in technique and organisation fell off sharply, either of which would lead to a declining rate of profit. The deceleration in production would naturally worsen the relative position of the entrepreneurs.

The suggestion that labour costs per unit of output were rising seems to be confirmed by the record of money wages and prices. It is not possible to show precise trends in product wage rates in manufacturing industry but use may be made of two substitute indexes. The trend growth rates are as follows.

PRODUCT WAGE RATES—AVERAGE ANNUAL GROWTH RATES

Period	(1)	(2)
1860–4 to 1870–4	+ 2·3%	+ 1·0%
1870–4 to 1893–7	+ 1·9%	+ 1·1%
1893–7 to 1909–13	− 0·3%	0·0%

(1) Money wage rate from [27] divided by Lewis's index of prices of manufactured goods [20].

(2) Money wage rate from [27] divided by Feinstein's index of prices of final products [9].

The two indexes disagree on the actual value of the trend but they agree broadly on the behaviour as between the trend periods. They agree that the sharp decline in labour productivity growth after the 1870's was not accompanied by a similar decline in growth of product wage rates and that after 1893–97 there was a sharp decline in the trend growth rate of product wage rates.[1] The series, therefore, suggest that profit rates per unit of output were declining from 1872 to 1895 and rising until 1913. Given the trends in output the well-known U-shaped trend of profits as a percentage of national income follows. The causal nexus is completed by the suggestions of Musson and Feinstein [24, pp. 211–2 and 9, pp. 370–2] that the trends in profits had repercussions on the level of investment through the factors of incentive and availability of finance.

D. J. Coppock

University of Manchester.

[1]The behaviour of money wage rates in the Great Depression presumably reflects the growth of Trade Union strength in the period in conjunction with a weak "market environment." See Phelps Brown and Hart [26].

REFERENCES

1. O. Aukrust, "Investment and Economic Growth," *Productivity Measurement Review*, 16 (1959), pp. 35-50.
2. Board of Trade, *British and Foreign Trade and Industry, Memoranda, Statistical Tables and Charts*, H.M.S.O., Cd. 1761, 1903.
3. T. H. Burnham and G. O. Hoskins, *Iron and Steel in Britain*, 1870-1930, Allen and Unwin, 1943.
4. A. K. Cairncross, *Home and Foreign Investment* 1870-1913, Cambridge University Press, 1953.
5. D. J. Coppock, "The Climacteric of the 1890's : A Critical Note," *Manchester School*, 1956, pp. 1-31.
6. ————————, "The Causes of Business Fluctuation," *Transactions of the Manchester Statistical Society*, 1959-60.
7. R. F. Crawford, "An Inquiry into Wheat Prices and Wheat Supply." *Journal of the Royal Statistical Society*, Vol. LVIII, 1895, pp. 75-111.
8. N. Deerr, *The History of Sugar*, Vol. 2, Chapman and Hall, 1950.
9. C. H. Feinstein, *Home and Foreign Investment, Some Aspects of Capital Formation, Finance and Income in the United Kingdom*, 1870-1913, Unpublished Ph.D. thesis, Cambridge University Library.
10. E. Frickey, *Production in The United States*, 1860-1914, Harvard University Press, 1947.
11. R. Goldsmith, "The Growth of Reproducible Wealth of the United States of America from 1805 to 1950," *Income and Wealth, Series II*, Bowes and Bowes, 1952.
12. A. Imlah, *Economic Elements in the Pax Britannica—Studies in British Foreign Trade in the Nineteenth Century*, Harvard University Press, 1958.
13. R. P. Higgonet, "Bank Deposits in the United Kingdom, 1870-1914," *Quarterly Journal of Economics*, Vol. LXXI, 1957, pp. 329-67.
14. W. Hoffman, *British Industry*, 1700-1950, (Translated by W. H. Chaloner and W. O. Henderson), Blackwell, 1955.
15. R. H. Hooker, "Farm Prices of Wheat and Maize in America, 1870-99,' *Journal of the Royal Statistical Society*, Vol. LXIII, 1900, pp. 648-57.
16. W. S. Jevons, *The Coal Question*, (Third Edition edited A. W. Flux), Macmillan, 1906.
17. C. P. Kindleberger, *The Terms of Trade, A European Case Study*, Technology Press of M.I.T. and Wiley, 1956.
18. W. T. Layton and G. Crowther, *An Introduction to the Study of Prices*, (Third Edn.) Macmillan, 1938.
19. League of Nations, *Industrialisation and Foreign Trade*, 1945.
20. W. A. Lewis, "World Production, Prices and Trade, 1870-1960." *Manchester School*, Vol. XX, 1952, pp 105-38.
21. ————————, and P. J. O'Leary, "Secular Swings in Production and Trade, 1870-1913, *Manchester School*, Vol. XXIII, 1955, pp. 113-152.
22. W. Malenbaum, *The World Wheat Economy*, 1885-1939, Harvard University Press, 1953.
23. M. G. Mulhall, *The Dictionary of Statistics*, (Fourth Edn.), Routledge, 1909.
24. A. E. Musson, "The Great Depression in Britain, 1873-1896 : a Reappraisal," *Journal of Economic History*, Vol. XIX, 1959, pp. 199-228.

25. E. M. Ojala, *Agriculture and Economic Progress*, Oxford University Press, 1952.
26. E. H. Phelps Brown and P. E. Hart, "The Share of Wages in National Income," *Economic Journal*, Vol. LXII, 1952, pp. 253-75.
27. ————————, with Sheila V. Hopkins, "The Course of Wage-Rates in Five Countries, 1860-1939, *Oxford Economic Papers*, New Series, Vol. II, 1950, pp. 226-96.
28. ————————, and S. A. Ozga, "Economic Growth and the Price Level." *Economic Journal*, Vol. LXV, 1955, pp. 1-18.
29. J. T. Phinney, "Gold Production and the Price Level : The Cassel Three per Cent Estimate," *Quarterly Journal of Economics*, Vol. XLVII, 1932-33, pp. 647-79.
30. A. R. Prest, "National Income of the United Kingdom, 1870-1946," *Economic Journal*, Vol. LVIII, 1948, pp. 31-62.
31. W. W. Rostow, *British Economy of the Nineteenth Century*, Oxford, 1948.
32. ————————, *The Process of Economic Growth*, (Second Edn.), Oxford, 1960.
33. W. Schlote, *British Overseas Trade from* 1700 *to the* 1930's. (Translated W. H. Chaloner and W. O. Henderson), Blackwell, 1952.
34. J. A. Schumpeter, *Business Cycles*, Vol. II, McGraw Hill, 1939.
35. C. Snyder, "New Measures of Trade and of Economic Growth," *Revue de l'Institut International de Statistique*, January, 1934, pp. 5-15.
36. ————————, "Growth of World Trade Vs. Basic Production," *Revue de l'Institut International de Statistique*, April, 1934, pp. 26-36.
37. B. Thomas, *Migration and Economic Growth*, N.I.E.S.R. Study XII, Cambridge, 1954.
38. M. J. Ulmer, *Trends and Cycles in Capital Formation by U.S. Railroads*, 1870-1950, (N.B.E.R. Occasional Paper 43), 1954.
39. U.S. Department of Commerce, *Statistical Abstract of the United States*, 1911.
40. ————————, *Historical Statistics of the United States*.
41. E. Wagemann, *Economic Rhythm*, (Translated by D. H. Blelloch), McGraw Hill, 1930.
42. Rosemary Walker, "A Note on Secular Swings around Alternative Trends," *Manchester School*, Vol. XXV, 1957, pp. 317-31.

The Rate of Interest and the Timing of Capital Projects

I

It is frequently stated by businessmen[1] and [2] that variations in the rate of interest have little or no influence on their investment decisions. Shackle, in an important article[3] gave the rationale for this view for the case in which interest rates have changed and are expected to remain at their new level permanently. Recent evidence[1] appears to suggest that businessmen hold the same view even in the case in which, as under credit squeeze conditions, interest rates can be expected to fall from their present high level in the next year or two. While it is probably true that the terms of borrowing do influence firms only slightly even under these conditions, the question needs to be asked whether it is economic wisdom for firms which are raising long term capital to behave in this way. Even in the simple situation in which interest rates are assumed to remain constant at the present level indefinitely, methods used by firms to calculate the rate of return offered by capital projects are often very crude and give results which differ widely both from each other and from accepted economic methods.[4] The credit squeeze situation is much more complex, because the whole structure of interest rates can be expected to move downwards in the future and the firm must consider the alternative means of raising capital for different long term periods, assess projects in this light, and calculate the advantages to be obtained from postponement and raising capital on less onerous terms. As far as the authors are aware, there is in fact no established economic technique for solving these problems. That the terms of borrowing in practice have little apparent effect, may, therefore, be no indication of their true economic importance.

[1]Radcliffe Report : Industrialists on the effects of changing interest rates. Vol. II, Part II.

[2]P. W. S. Andrews, *Oxford Economic Papers*, Vol. 3, 1939.

[3]G. L. S. Shackle, " Interest Rates and the Pace of Investment," *Economic Journal*, March 1946.

[4]Anthony Merrett and Allen Sykes, " Calculating the Return on Capital Projects," *Journal of Industrial Economics*, October 1960.

The purpose of this article is to develop the technique required to evaluate capital projects under credit squeeze conditions, and by calculating the numerical effect of changing interest rates, assessing the extent to which interest rates, correctly considered, could cause postponement of investment expenditure.

II

It may be worthwhile to begin by pointing out that this is a real problem, the answer to which is by no means obvious. Suppose, for example, that the firm can currently borrow on debenture terms at the rate r_1, but, as is commonly the case under credit squeeze conditions the firm anticipate with confidence that from the end of, say, year k it will be able to raise debenture capital at the rate r_2, where r_2 may be considered the *lower* long term rate of interest. Consider first the criterion required to decide whether or not the project offers any net present value if commenced immediately and financed by borrowing on debentures at the rate r_1 for p years.

The conventional present value criterion is that

$$R = \sum_{i=1}^{i=n} \frac{A_i}{(1+r)^i} - C > 0 \qquad (1)$$

where R is the net present value, A_i is the net cash flow[1] assumed to be arising end year i, C is the initial capital cost and r is the rate of interest representing the marginal cost of capital. It should be evident that the criterion given in (1) with $r = r_1$ cannot apply to the case now being considered. The criterion given in (1) makes no mention of the period for which the loan required for this project is to be undertaken. It is evident, however, that if the life of the loan extends beyond the life of the project this has a serious bearing on the profitability of the investment. Many projects which would be acceptable if the capital could be borrowed for their life only would become unprofitable if circumstances were such that capital had to be raised for a longer term at the higher interest rates. Criterion (1) cannot therefore be generally valid when the loan extends beyond the life of the project. We shall also demonstrate later that it is not even valid, under the conditions

[1]The net cash flow is defined as profit less tax when actually paid, plus depreciation less capital expenditure when actually paid.

assumed here when the life of the loan and the project coincide, if this period extends beyond the date at which interest rates are expected to change.

Again it should be apparent that if the firm is planning to raise long term capital at the beginning of the year j ($j \leqslant k$) at the high rates then obtaining and plans to be partially self-financing, the sums arising from a project before the year j are particularly advantageous in that they enable the firm to avoid raising long-term capital at these abnormally high rates. This fact should be reflected in the method of discounting the revenues promised by the project, and it clearly is not in equation (1). Finally it should be noted that the general problem under the conditions postulated is not merely acceptance or rejection of a project but also, if the project is accepted, to ascertain the most advantageous time to raise capital and commence it. Postponing commencement may mean that the firm can avoid raising long term capital at the current high rates, but it also involves the firm postponing receipt of the net revenues from a project to a later date: these two factors must be appropriately evaluated against each other.

III

To construct a more general theory which will enable us to solve these problems, we first of all need to construct a more general method reducing future outlays to their present value. The particular case which we will consider is the fairly typical one in which the firm obtains its finance during the period of high interest rates from several sources by drawing on its own profits, bank overdrafts, long term borrowing and possibly by raising equity capital. We shall assume, as above, that the firm anticipates that the long term rate of interest will fall from its present high level, r_1, to its normal *lower* level, r_2, at the beginning of year $k+1$ and to remain constant thereafter.[1] We shall also assume that the firm has such investment opportunities relative to its own financial resources that it can confidently expect to have a net cash outflow—i.e. to be a net borrower—up to and including year k. Throughout the following analysis we shall assume that the profitability of individual investment projects is entirely independent of any other project being accepted or rejected.

[1]Throughout this paper it is assumed that the optimal means of finance has already been determined. The authors hope to consider this problem, which is too large to be dealt with here, in a subsequent paper.

Consider first the simple situation in which the firm raises the capital it requires on a conventional bank overdraft at the rate r_1, which reduces at the beginning of year $k+1$ to the rate r_2 (the overdraft rate is assumed to fall with the general structure of interest rates at the end of the credit squeeze). Under these circumstances the net present value of the project is :

$$R = \sum_{i=1}^{i=k} \frac{A_i}{(1+r_1)^i} + \sum_{i=k+1}^{i=n} \frac{A_i}{(1+r_1)^k(1+r_2)^{i-k}} \qquad (2)$$

Identically the same equation would hold if the firm, as is more probable, planned to raise capital on bank overdraft until the period of lower long term interest rates began, and then to repay the overdraft by raising long term capital. Suppose now that a firm, having raised temporary finance by borrowing on a bank overdraft anticipates that it will be obliged to raise debenture capital during the current period of high interest rates for a period of p years. For simplicity we shall assume that the firm is considering raising the long term capital now at $t = o$. (With minor modifications the following analysis can be adapted to cover the more general situation where the firm expects to raise the capital at some future date $j \leqslant k$). Suppose also, owing to flotation expenses, etc., it decides, as is commonly the case to raise sufficient capital at this time to cover both its current requirements and those for the period up to the end of year k.

The formulas we shall now develop will be based on calculating the replacement cost of postponing the acquisition of funds from one year to another. The rate of discount required to bring sums arising after the year k to their value at the *beginning of year* $k+1$, is clearly the rate r_2. For after this date the firm can borrow at the rate r_2 to replace funds,[1] the acquisition of which has been postponed, and the firm will be indifferent, therefore, between, say, the sum $A/(1+r_2)^m$ at the beginning of year $k+1$ and the sum A at the end of year $k+m$.

[1]To cover the possibility that the firm after year k may be forced to borrow for a fixed period, it will be assumed that the firm can always invest funds at the rate r_2. It may arise that the sums recovered from the project in any one year cannot be used to repay the original debt and cannot be re-invested at r_2 or more. The method developed below can be adapted to cover this case, but for simplicity of exposition we will assume that the firm can re-invest at r_2.

It is now necessary to consider the rate which should be used to bring the sums arising in years 1 to k inclusive to their present value at $t = o$. We have assumed that the firm will be raising capital for the years o to k at $t = o$. The rate of interest at which the firm would be willing to postpone acquisition of funds within the interval of time o to k inclusive depends on the use to which it can put the capital it has raised at time $t = o$. If the firm could invest the capital raised short term until required for its long term investments only at the rate r_m, then this is the rate of discount which should be applied in this time interval. Since the actual magnitude of this rate is not of appreciable importance in our later calculations we will for simplicity of notation take this rate also as r_1.[1]

We have so far established the rates of discount that will enable us to calculate the sums at certain different times between which the firm would be indifferent. The evident gap in the analysis is the appropriate discounting to be applied to establish the equivalent sums in the years up to and including year k to other sums arising *after* year k. For the sums arising at or before the end of year k have the important advantage over sums arising later in that they can be taken into account and in part used to finance the capital programme of the years up to k, and thus avoid borrowing at the high rate of interest r_1. (In so far as some of these funds are distributed as profits as well as being used for self-financing, they must be supposed to yield the same return to the firm at the margin).

Consider the replacement cost to the firm of £1 at the end of year k. This is evidently the interest burden the firm would have to assume in order to replace the £1. This sum would be replaced by borrowing £1 more at $t = o$. On our assumptions these sums raised can be re-invested to cover their interest charges up to the end of year k. The burden incurred by raising the additional capital is therefore the interest r_1 for the remaining years of the loan, that is from k to p, and interest r_2 for all the years thereafter. Discounting at the rate r_2 this has a present value *at the beginning of year* $k+1$ of :

[1]Where the sums raised in advance cannot be re-invested short term to cover the interest burden they involve, then this represents a debit against the projects for which this extra capital has been raised. This point is easily covered by the method developed below, but again for simplicity of exposition, it will be ignored.

$$v = \sum_{i=1}^{i=p-k} \frac{r_1}{(1+r_2)^i} + \sum_{i=p-k+1}^{i=\infty} \frac{r_2}{(1+r_2)^i} \tag{3}$$

By elementary reduction (3) becomes :

$$v = \frac{r_1}{r_2}[1-(1+r_2)^{-p+k}]+(1+r_2)^{-p+k} \tag{4}$$

Now the firm must be indifferent between (i.e. value equally) £1 at the end of year k and v at the beginning of $k+1$. For if the firm valued v at the beginning of year $k+1$ at more than £1 at the end of year k it would reduce the amount of capital raised for that year until this inequality no longer held. Similarly the firm cannot value the £1 at more than v or it would raise more capital until this inequality no longer held—that is until the value the last additional £1 raised and available in year k is exactly equal in value to the sum v at the beginning of year $k+1$. Hence to find the value at end of year k of any sums arising after year k we first bring them to present value at the beginning of year $k+1$ and divide by v. To state this point another way, £1 at the beginning of year $k+1$ is equal in value to $1/v$ at the end of year k. Re-arranging $1/v$ we obtain :

$$\frac{1}{v} = \frac{1}{\frac{r_1}{r_2} + \frac{1-\frac{r_1}{r_2}}{(1+r_2)^{p-k}}} \tag{5}$$

This rather clumsy form of expression will be found useful for our later discussion.

The case in which the capital is raised on equity terms can be looked at as one in which the firm intends to maintain the rate of return r_1 on the capital raised as a matter of policy, and is the special case in which $p \to \infty$ and (5) becomes simply $\frac{r_2}{r_1}$. This result is in accord with commonsense : where the firm is being obliged to raise capital at the rate of interest r_1, the sums arising before the beginning of year k, in effect offer an interest saving of r_1 into perpetuity, compared with r_2 into perpetuity offered by the sums arising after this date. The discounted value of sums arising

immediately after the end of year k are, therefore, worth only $\frac{r_2}{r_1}$ of the sums arising immediately before that date.

Precisely the same form of discounting as has been developed for the net cash flows from the project must be applied to the interest obligations and the terminal repayments of loan capital, incurred to finance the project under consideration. Thus the net present value offered by the project is :

$$R = \sum_{i=1}^{i=k} \frac{A_i}{(1+r_1)^i} + \frac{1}{v} \sum_{i=k+1}^{i=n} \frac{A_i}{(1+r_1)^k(1+r_2)^{i-k}} - L \qquad (6)$$

where L is the net present value of the interest and capital repayment obligations given by :

$$L = \sum_{i=1}^{i=k} \frac{Cr_1}{(1+r_1)^i} + \frac{1}{v} \sum_{i=1}^{i=p-k} \frac{Cr_1}{(1+r_1)^k(1+r_2)^i} + \frac{C}{v(1+r_1)^k(1+r_2)^{p-k}} \qquad (7)$$

$$= C[1-(1+r_1)^{-k}] + C(1+r_1)^{-k} = C$$

Thus we reach the important conclusion that it is still permissible to use the total capital outlay in the net present value calculation.

The following characteristics of (6) and (7) should be noted. First, the generalized present value given in (6) reduces to ordinary present value (1) as a special case. This follows immediately from the fact that $\frac{1}{v} = 1$ when $r_1 = r_2$. Second, in the case in which the life of the project extends beyond the life of the loan p ($p > k$) the result given in (7) still holds, if the firm can obtain further capital at the rate r_2. For in this case when the firm renews the loan in year m, say, the present value of the new set of interest (at rate r_2) and capital repayment obligations, discounted at rate r_2 back to year m is simply C—precisely what it would be if the firm simply repaid the existing loan in that year.

Finally we may note that (6) reduces to a particularly simple and illuminating form in the special case of a capital project with constant net cash flow A per annum, a *perpetual* life span, and which is being financed by raising equity capital. In this case p tends to infinity and $\frac{1}{v}$ tends to $\frac{r_2}{r_1}$. The net present value offered by the project prior to the credit squeeze when capital could be

raised at the rate r_2 was $\frac{A}{r_2}-C$. The net present value during the credit squeeze as given by (6) is simply $\frac{A}{r_1}-C$.

IV

Investment Programming

For the numerical analysis given in the following section we need to establish some basic principles concerning the optimal time at which to raise capital and commence the capital projects. This is not essential for an understanding of the numerical analysis and may be omitted by readers not interested in the technicalities of the method.

Consider the effect of postponing the commencement of the project from the present time to the beginning of year t where $0<t\leqslant k$. This postponement shifts all the A_i's t periods along the time axis, and this results in the following additional discounting of the net cash flows.[1]

Consider the additional discounting of the net cash flow A_i and let $m = k-i\geqslant 0$. Where $t>m$ then the postponement takes the net cash flow outside the credit squeeze period and results in it being multiplied by the *additional* factor $\frac{1}{v(1+r_1)^m(1+r_2)^{t-m}}$. Where $t<m$ the factor $\frac{1}{v(1+r_2)^{t-m}}$ is omitted from this additional discounting which becomes $\frac{1}{(1+r_1)^m}$. If $m<0$, the additional discounting is simply $\frac{1}{(1+r_2)^t}$. It is seen, therefore, that postponement results in different additional discounting of the net cash flows depending on their place in the sequence.

Consider now the sequence of net cash flows A_i-Cr_1 ending in year n (the last year of the project's life) with the net cash flow $A_n-C(1+r_1)$. (The assumed repayment of capital at the end of the project's life, for the reasons given in the penultimate paragraph of the last section, involves no loss of generality). Assume this sequence of net cash flows (with the exception of the last) are all positive or zero.

[1] In what follows we are assuming the net cash flows and capital outlays are unchanged by postponement.

From the discussion given above, postponement will result in these net cash flows being multiplied by additional factors *at most* equal to $\frac{1}{(1+r_2)^t}$ and the earlier net cash flows by factors less than this. On the other hand the final (probably negative) net cash flow is multiplied by exactly $\frac{1}{(1+r_2)^t}$ owing to the postponement. The ultimate result, therefore, of postponement within the period 0 to $k+1$ must reduce net present value offered by the project. Similarly postponement after the beginning of year $k+1$ when the advantages of the lower interest rates have already been secured also reduces net present value. Hence the optimal time of commencement is either 0 or the beginning of year $k+1$. The net present value from the latter date of commencement will be the R of equation (1) (with $r = r_2$) multiplied by $\frac{1}{v(1+r_1)^k}$.

These results need not hold if one or more of the terms assumed positive or zero above is negative. For in this case the differential discounting of the net cash flows may result in a higher net present value if the project is commenced in a year between 0 and $k+1$. Thus a project requiring additional capital outlays in year k of its life ($A_k<0$) may offer a higher present value if postponed one year so that the additional capital outlay of year k can be financed at the more favourable terms which obtain after that year. In the following numerical analysis the conditions of the previous paragraph are assumed to obtain.

V

We shall now consider the numerical importance of the theory outlined above, in terms of the range of projects significantly affected by the credit squeeze conditions and for which postponement is the more profitable alternative. By range of projects is meant projects with differing economic lives and internal rates of return (see footnote[4], p. 233). We consider first projects having a fixed economic life, constant net cash flows and not subject to renewal at the end of the life of the initial project.

Table 1 sets out the effect of the changing interest rates when the projects are financed by equity or debenture capital. The zero present value rates are defined as the rates of return on projects

at which the present value would be reduced to zero by the change in interest rates, if immediate commencement is proposed. These rates indicate how the credit squeeze which automatically reduces the attractiveness of *all* impending investments, operates on those which are marginal.

The breakeven present value rates define the rates of return which projects must offer or exceed before, with the changed interest rates, it is more profitable to undertake immediate commencement than to postpone. By " more profitable " is meant " offers a higher present value."

TABLE 1

Life of Project in Years	Zero Present Value Rates of Return	
	Equity	Debentures
	per cent	per cent
10	7·9	3·7
20	6·5	3·3
30	5·9	3·0
40	5·6	2·9
Life of Project in Years	**Break Even Rates of Return**	
	Equity	Debentures
	per cent	per cent
10	30·0	18·3
20	23·0	13·0
30	18·3	10·3
40	17·7	9·0

For *net of tax* rates of raising capital :

Equity : $r_1 = 5$ per cent, $r_2 = 4$ per cent

Debentures : $r_1 = 3$ per cent, $r_2 = 2\frac{1}{2}$ per cent

The term of the debenture is $p = 22$, and the remaining duration of the credit squeeze is $k = 2$.

The projects given in the table are assumed to have a series of constant net cash flows and the calculation of the rates of return to be net of tax.[1]

It is seen from Table 1 that if projects offer a very high rate of return it is more profitable to undertake immediate commencement and finance, regardless of the higher cost of financing. This again is in conformity with commonsense since if the net present values are large relative to capital cost, it is evident that the cost of financing the project is relatively unimportant compared with the cost of the postponement of the very high present value to be obtained from the project. Also, for a given rate of return, immediate commencement is more attractive than postponement, the longer the life of the project. This is due in part to the fact (see Table 2) that for a given rate of return, net present value is higher the longer the project's life. As might be expected, both the zero present value rates and the break even rates are decidedly less for debenture case than for the equity.

Before the full significance of these rates can be assessed we must first deal with the problem of risk. In so far as the rates of return calculated on a project are subject to risk, we may as an approximation, think of the nominal rate of return offered by the project as being subjectively reduced, by the appraisal of the firm, to a lower certainty equivalent. If this is accepted then the results of Table 1 hold, but applied to these *certainty equivalent rates of return.*

It is seen that for both forms of finance the break even rates are high and it is unlikely that many projects of this type can offer *certainty equivalent* rates of return of these magnitudes. Also the zero present value rates are quite high and for projects to offer a commercially acceptable return over the cost of borrowing these rates would have to be considerably exceeded. This indicates that many projects are likely to be postponed simply on the grounds that the net present value to be obtained from them *under existing conditions is insufficient* even though such postponement offers a lower net present value *now*. The reason for this slightly

[1]To this extent the calculation as given here is not quite realistic since with new ventures (see below) investment allowances in practice can make the initial years more profitable. This, however, is partially offset by the only gradual build up of revenue and we have therefore chosen to assume constant revenues throughout the project's life.

paradoxical situation is simply that, although insufficiently profitable for immediate commencement, the net present values offered by the projects once the credit squeeze conditions have been lifted will revert then to the sufficiently profitable level which obtained before the credit squeeze.

In the light of these comments it may be helpful to demonstrate the effect on net present values of changes in interest rates and the gains offered by postponement, for projects with, what may be supposed to be, common certainty equivalent rates of return. This is given in Table 2.

TABLE 2

EFFECT ON NET PRESENT VALUE OF CHANGING INTEREST RATES

Project with IRR of 10 per cent Financed by Equity Capital	Life of			
	10 years	20 years	30 years	40 years
$r_1 = r_2 = \cdot 04$, gives p.v. of ...	320	597	835	1025
$r_1 = \cdot 05$, $r_2 = \cdot 04$, $k = 2$ immediate commencement offers p.v. of	97	298	480	628
Postponed 2 years offers p.v. now of	232	433	606	744
Project with IRR of 7 per cent Financed by Debenture Capital	**Life of**			
	10 years	20 years	30 years	40 years
$r_1 = r_2 = \cdot 025$, gives p.v. of ...	246	472	687	883
$r_1 = \cdot 03$, $r_2 = \cdot 025$, $p = 22$, $k = 2$ immediate commencement offers p.v. of	170	372	569	749
Postponed 2 years offers p.v. now of	216	413	601	772

Initial capital cost taken as 1000 throughout.

The upper part of the table gives in the first row the net present values, pre-credit squeeze, of projects each with an internal rate of return of 10 per cent but with different economic lives. The second row gives the net present values resulting from these projects with the given increase in long terms interest rates

and with immediate commencement. The third row gives the net present value *now* which would result from postponement of the projects for two years and financing them by equity capital at the lower rates then ruling. The lower part of the table gives similar results for the case of projects with an internal rate of return of 7 per cent financed by raising debenture capital for twenty-two years. The impact on net present value is seen to be substantial in the equity case and rather less in the case of debentures.

The results given in the above tables are for specific parameters and interest rates. It will be found, however, that within the range of value that can reasonably be expected to obtain, the conclusions suggested by the tables hold generally for the type of projects to which they refer.

These simple types of capital projects cannot, of course, be taken without modification as a guide to the complex and diverse projects which commonly arise in business. These results offer some guidance as to orders of magnitude, however, if we distinguish between the following types of investment project.

Replacement Investment: This is capital expenditure conditional on the termination of the economic life of an existing asset ; and is typically a form of cost saving investment.

Expansion Investment: This is investment in the firm's existing line of business and for which it already has some production and distribution capacity.

New Ventures: These are investments which are additional to the firm's existing line of business in the sense that it has no existing capacity which could be used as an alternative to immediate investment in these ventures.

If the firm exercised proper care in reviewing its investment possibilities it would normally (where the marginal cost of capital remains constant) undertake the first two categories of investment in particular assets, as soon as the return from immediate commencement offered sufficient compensation (compared with later dates of commencement). Although these projects may offer very high rates of return over their *whole* lives, the *additional* gain from commencing in one year rather than another may be quite small. This is due to the fact that in this type of investment,

profitability tends to build up over time, until a point is reached at which it is at last worthwhile to make an immediate commencement. Thus in the case of replacement expenditure, its profitability compared with the alternative of continuing with the existing assets, builds up with the decline in efficiency of existing assets due to age and wear and tear, and the increase in the economic efficiency of new replacements with normal technological progress. Similarly expansion investment becomes increasingly attractive with increasing demand for the product and the high costs of operating existing equipment at high levels of output. Again at a given point it becomes marginally more profitable to undertake expansion rather than postpone for another year.

It would appear from the discussion given above that some of these projects, despite their long economic lives (taking into account that of their replacements) would probably be more profitable if postponed, if they are to be financed under credit squeeze conditions by raising long term capital. (This, however, is open to, and should be the subject of, empirical verification).

The results of Tables 1 and 2 apply more directly to the case of new investment ventures. In many instances these are not investment possibilities which the firm has been waiting to mature, but new possibilities only recently discovered and which may well be highly profitable in the early years. A qualification to this, however, is that the firm may anticipate that if the initial project is successful then at the end of its economic life a similar replacement project may be commenced. It is commonly the case that these continuations of the original project are not taken into account when evaluating the initial project. This is perhaps justifiable with long lived projects or where for other reasons great uncertainty surrounds the later life of the project or the firm's willingness at a later date to undertake a continuation.

If with a relatively short lived project the successive continuations can reasonably be taken into account and these have the same certainty equivalent rate of return as the initiating project, the effect as might be expected is to make the immediate commencement of the project considerably more attractive. The break even rates for this case are given in Table 3.

TABLE 3

BREAK EVEN RATES FOR PERPETUALLY RENEWABLE PROJECTS

Financed by	Economic Life of Successive Renewals	
	10 years	20 years
	per cent	per cent
Equity	15·6	15·6
Debentures	10·4	7·5

Interest rates and other parameters as for Table 1.

In this paper we have attempted to discuss the technique as briefly as is consistent with demonstrating its validity. The removal of some of the assumptions on which the technique is based, while involving modification of the techniques, involves no substantial alteration of the general results obtained.

Regarding the general application of the results obtained, however, we would not wish to minimize the fact that we have assumed a higher standard of financial analysis of investments than is commonly practised in British industry. In particular replacement investments are frequently overlooked or deliberately postponed until the gain from immediate commencement is high. Also expansion investment is sometimes undertaken on grounds of long term profitability without regard to the higher profitability of alternative dates of commencement.

Conclusions

The main conclusions suggested by this article are as follows :

(1) The technique required to calculate net present value when interest rates are expected to fall differs in important respects from that required when interest rates remain constant. Where the latter is applied to credit squeeze conditions it would appear likely to lead to serious errors.

(2) There is a reasonably strong *prima facie* case that investment decisions should be quite sensitive to changes in the

long term rates of interest of the magnitude experienced in post-war Britain, if (*a*) the firm is forced to borrow long term, (*b*) the revenues and capital cost are not significantly altered by postponement, and (*c*) the correct method of analysis is used.

(3) The commonly argued view that the magnitude of the changes in interest rates is dwarfed by the risk allowances does not hold in this case.

ANTHONY MERRETT and ALLEN SYKES

London.

The Cost Function for Crude Oil Production

Introduction

In this paper is presented a cost function for crude oil production which contains the major factors affecting the cost of getting crude oil from the reservoir to the top of the well.

The factors considered to be of major importance are: (1) drilling costs, (2) well operating costs, and (3) costs of physical waste and depletion.

Consequently, the cost function for crude oil production can generally be described as follows: Crude oil costs of production $= F$ (drilling, well operating, physical waste, depletion).

The particular technology of crude oil production reveals the fact that the production rate with which the crude oil of a reservoir is produced, under certain geological conditions, determines average costs (the cost per barrel) as well as the ultimate recovery rate, if a certain production rate is continued over the lifetime of the pool. By increasing the production rate, it is possible to decrease the cost per barrel, but on the other hand, because of the special production mechanisms of the reservoirs, an increase of the production rate will lower the ultimate recovery rate.

To allow for depletion and physical waste the factor q_r is introduced, representing the ratio of the exploitation or ultimate recovery if a reservoir is depleted at a certain production rate per period to the original total supply of crude oil in place.

The hypothesis concerning the relationship between q_r and the production rate (r) is that q_r decreases at an increasing rate as the production rate increases.

On the basis of empirical data the particular nature of $q_r = f(r)$ has been investigated, and it was found that the hypothesis concerning this relationship is empirically valid.

(A) Drilling Costs

In 1954 a joint inquiry was held by the " American Petroleum Institute," the " Independent Petroleum Association of America," and the " Mid Continent Oil and Gas Association " in order to determine the relationship between average drilling costs per well and depth of the wells drilled. The result of this important survey is shown in the following table :

U.S. AVERAGE DRILLING COST PER WELL AND PER FOOT IN 1953

Depth Range	Average Depth	Average Cost	Average Cost	Wells Drilled	Total Footage
	per well	per well	per foot	U.S.	000 feet
0— 1,250	849	$5,600	$6·63	5,001	4,244
1,251— 2,500	1,793	$14,300	$7·99	10,219	18,323
2,501— 3,750	3,121	$28,400	$9·10	12,603	39,336
3,751— 5,000	4,292	$43,300	$10·09	7,442	31,944
5,001— 7,500	6,104	$76,400	$12·51	8,624	52,638
7,501—10,000	8,631	$128,300	$14·86	3,496	30,173
10,001—12,500	11,083	$230,800	$20·83	1,414	15,671
12,501—15,000	13,380	$363,300	$27·15	447	5,981
over 15,000	16,038	$644,800	$40·20	33	529

Source : *World Oil*, Vol. 142, p. 138, February 15, 1956 and *The Oil and Gas Journal*, Vol. 54, p. 124, January 30, 1956.

In using the figures shown in the table above to determine a functional relationship between average drilling cost per well and depth, it seems that a quadratic function fits the data best :

$$C_D = B_0 + B_1 d + B_2 d^2$$

C_D = Average drilling cost per well, in tens of thousands of dollars.

d = Average depth per well, in thousands of feet.

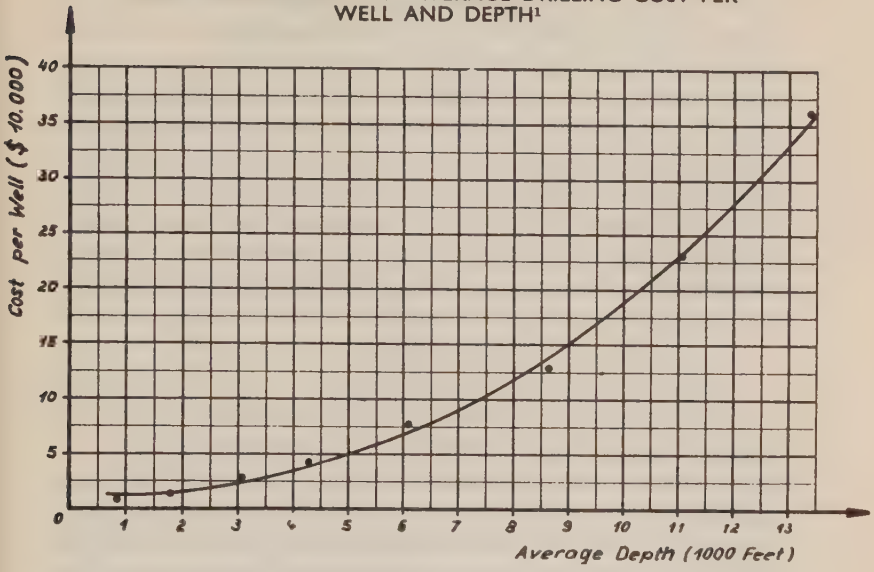
The following coefficients were then obtained by the use of the least squares method :

$$C_D = 1{\cdot}57 - 0{\cdot}50586d + 0{\cdot}22674d^2$$

The empirical data and the curve fitted to these data are shown in Graph 1 :

GRAPH 1

RELATIONSHIP BETWEEN AVERAGE DRILLING COST PER WELL AND DEPTH[1]



(B) Well Operating Costs

For the purpose of determining an empirical cost function showing the relationship between well operating costs of a certain pool and output (production rate), it was impossible to obtain meaningful statistical data. This particular problem, therefore, can only be discussed theoretically. The results of empirical investigations in other industries, however, do justify the conclusion that it is quite reasonable to assume linear cost–output relationships within certain limits.[2]

[1] In Graph 1, as well as in computing the quadratic function, depths of over 15,000 feet were eliminated because of their relative rarity.

[2] Concerning the validity of the assumption of linear cost–output relationships see especially :

(*a*) J. Johnston : " Statistical Cost Analysis," McGraw-Hill Book Company, New York, Toronto, London, 1960, chapter 4 where the author presents his own studies in this field, and chapter 5 where he discusses other important empirical studies.

Summarizing the empirical evidence, the author writes : " . . . the various short-run studies more often than not indicate constant marginal cost and declining average cost as the pattern that best seems to describe the data that have been analyzed," p. 168.

(*b*) Reiner Reustle : Ueber Untersuchungen zur Feststellung der empirischen Kostenfunktion, Diss., Freiburg/Brsg., 1958.

The problem of cost behaviour in relation to output is to be discussed from a short-run and long-run point of view :

(*a*) Short-run function : A cost function of this kind indicates cost–output relationships under short-run conditions, in this case mainly under the condition that the number of producing wells in a particular area remains constant. Changes in output per period are, therefore, due to changes in the speed of pumping, i.e. to changes of the production rates per period of the given wells. The assumption of short-run conditions in this case is equivalent to the assumption of a given reservoir already explored and developed. The cost–output relationship can thus be expressed by the equation :

$$C_1 = a_0 + a_1 r_1$$

C_1 = Total well operating costs per period.

r_1 = Production rate ; barrels per period holding number of producing wells in a particular area constant.

The coefficients a_0 and a_1 indicate the fixed and variable portion of total costs. The question of including exploration and development costs of a pool as well as drilling costs of the existing wells depends on the methods used in writing off these items. For instance, if constant rates per accounting period are practised, the corresponding amounts should be included into a_0. If, on the other hand, the amount to be written off is connected in some way with the production rate, this item has variable character and hence appears in a_1.

(*b*) Long-run function : The conditions of a long-run approach in our case are present if the number of producing wells of a certain reservoir varies (increases). Consequently, a long-run cost output function would be a function showing the relationship between cost and changes in output per period due to increased number of producing wells. This assumption is equivalent to the assumption of an explored but not yet fully developed reservoir. A long-run cost function can be expressed by the following equation :

$$C_2 = a_0 + a_1 r_2$$

C_2 = Total well operating costs per period.

r_2 = Production rate ; barrels per period, varying number of producing wells in a particular area ; output per well is held constant.

The coefficients a_0 and a_1 include again the fixed and variable portions of total costs.

From the long-run point of view, also, the treatment of exploration and development costs and the costs of the wells already drilled is in principle a question of the methods used in writing off these costs. In other words, the location and the shape of an empirical cost curve in relation to output depends, because of the great amounts of these items, upon writing off methods. It seems to be quite reasonable, however, to assume that if the items in question are partly included in the first type or fully in the second type of cost function, the corresponding curves appear to have a comparatively high portion of fixed costs and a comparatively low portion of variable costs.[3]

(C) Physical Waste and Depletion

The calculation of average costs (the cost of a barrel of crude oil produced) commonly includes the allowance of a certain percentage to cover depletion and waste costs. These percentages are customarily determined by present tax legislation. However, in the following paragraphs a method is advanced for incorporating these allowances into an empirical cost function according to the factual situation of a given reservoir. Furthermore, on the basis of empirical data, the kind of a relationship that exists between physical waste and some important factors influencing physical waste will be shown.

(*a*) Production rate and exploitation rate: Under certain geological conditions of a reservoir and with a given production mechanism (drive), the exploitation rate (or rate of ultimate recovery) mainly depends upon the production rate or the speed of producing the crude oil. For each individual production rate or production rates within certain limits, there exists a particular ratio (q_r) between exploitation and total supply of a pool under given geological conditions.

$$q_r = \frac{E}{S}$$

q_r = Efficiency rate for a given r.

[3]The problems of adjustments and statistical estimation of cost–output relationships are described in J. Johnston, *ibid.*, Chapter 3, pp. 26—43.

E = Exploitation or ultimate recovery, i.e. that amount of crude oil which can be produced from the total supply if a reservoir is operated at a certain production rate.

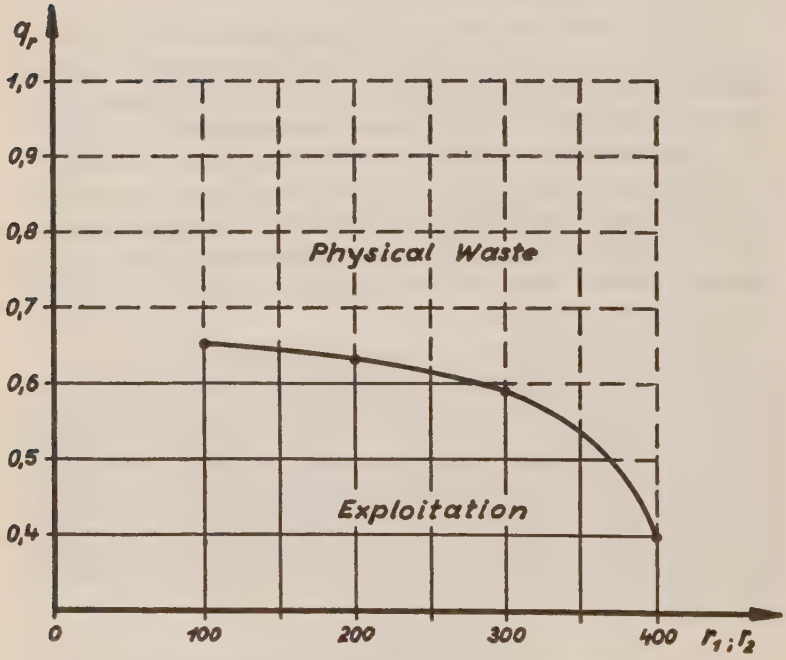
S = Original total supply of crude oil in place.

Because of the particular conditions of crude oil production, i.e. the influence of a certain drive mechanism, it is impossible to produce the total supply of a reservoir. Therefore, q_r will always have a value between 0 and 1.

The relationship between q_r and alternative production rates (r) can be described by the following curve :

GRAPH 2

HYPOTHETICAL RELATIONSHIP BETWEEN q_r AND ALTERNATIVE PRODUCTION RATES (r) [4]



[4] The abscissa is labelled r_1, r_2 indicating the hypothesis that the effect of alternative production rates on ultimate recovery (E) and hence on q_r is basically the same whether achieved by increased production rates due to increased speed of pumping (short-run viewpoint) or increasing the number of wells on a certain field (long-run viewpoint).

The curve $q_r = f(r)$ described in Graph 2 is definitely downward sloped, since an increase of the production rate (r), as previously mentioned, causes a decrease of the efficiency rate or an increase of physical waste.[5] Concerning an assertion about the shape of this curve, it should be expected to be convexly shaped. In other words, it is assumed that switching over from a production rate of 100 units per period to a production rate of 200 units per period would cause a smaller decrease in the exploitation rate (or a smaller increase in the waste rate) than switching over from 300 to 400 units per period.

It should be emphasized that a special function $q_r = f(r)$ is only valid under certain geological conditions. Concerning these conditions the reservoirs are usually distinguished on the basis of their drives. According to the gas–oil ratio present in a reservoir three types of drives are predominant :

1. Gas cap drive.
2. Dissolved gas drive
3. Water drive.

These three types of drive affect costs of well operation as well as the q_r function.

In connection with the q_r function the possible effect of modern production methods has to be mentioned. In order to increase ultimate recovery the gas or water pressure of the reservoirs is sometimes artificially maintained by gas or water injections. Their effects on the q_r function can be described in the following way :

(*a*) Production without any injections : In this case the relationship between q_r and r is represented by a curve like the one shown in Graph 2.

(*b*) Production with a single gas or water injection : When the injection becomes effective the performance of a single gas or water injection will shift the q_r curve to a higher level. The shape of the curve might also probably be flatter.

[5]The curve describes the relationship between q_r and r according to the particular technology of crude oil production, i.e. the higher the production rate per period the less will be the ultimate recovery because of the gas pressure decline or oil shrinkage (gas cap—or dissolved gas drive) or water pressure decline (water drive).

(*c*) The continuous performance of pressure maintenance can be expected to shift the curve, compared with cases (*a*) and (*b*), to a higher level, and probably to flatten the curve even more.

Between the factors

S = Original total supply of crude oil in place.
r = Production rate per period.
q_r = Efficiency rate for a given r.
t = Lifetime of the pool.

there exist the following relationships :

Expressed in terms of total physical waste :

$$\frac{(1-q_r)S}{t} = \frac{r}{q_r} - r,$$

where $(1-q_r)S$ = total physical waste.

Or expressed in terms of total supply :

$$\frac{S}{t} = \frac{r}{q_r}$$

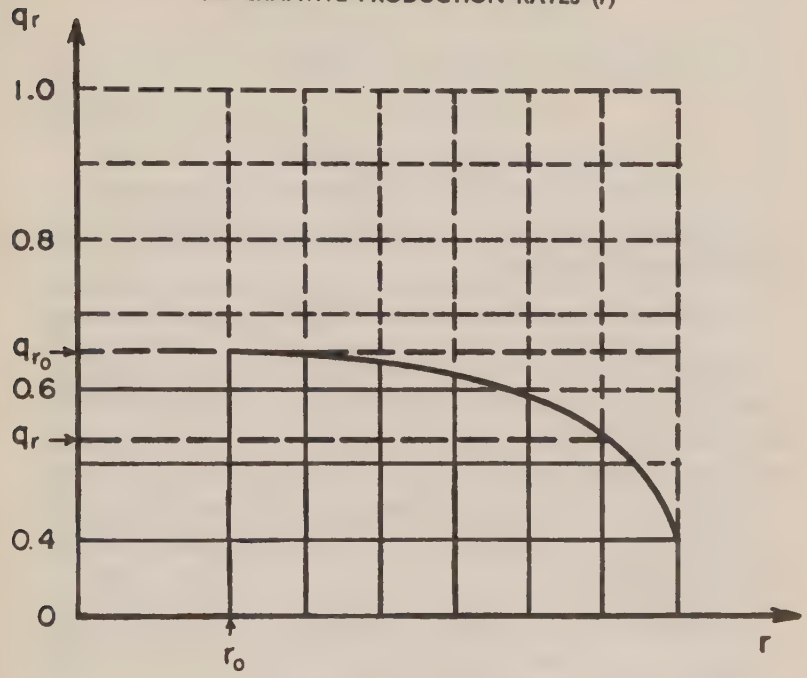
The quotient $\frac{r}{q_r}$ indicates current depletion and its associated physical waste under the given production rate per period and the other conditions of the pool. The formula as given, however, is only correct under the (invalid) assumption that the optimum efficiency rate attainable is equal to one ($q_{r_0} = 1$), i.e. that we can consider the difference between S and E_r (exploitation or ultimate recovery at a given r) as physical waste ; or in particular that :

$$S - E_r = S - Sq_r = \text{total physical waste.}$$

Since, under the given conditions of the pool (and no improvements by any injections—see example shown in Graph 2) the optimum efficiency rate is about 0·65 ($q_{r_0} = 0{\cdot}65$), physical waste has to be measured relative to that q_{r_0}. This aspect still holds if improvements are made by injections, because it simply means that we can consider as waste only that portion which is the difference between the optimum exploitation at the optimum production rate (E_0 at r_0) and the exploitation attainable at any other production rate (E_r at r).

GRAPH 3

HYPOTHETICAL RELATIONSHIP BETWEEN q_r AND ALTERNATIVE PRODUCTION RATES (r)



The relationships between the relevant variables can now be expressed as follows :

$$E_0 - E_r = \text{actual physical waste.}$$

$$\frac{E_0 - E_r}{S} = q_{r_0} - q_r = q_{r_w} \text{ (actual waste rate).}$$

$$\frac{E_0 - E_r}{t} + \frac{E_r}{t} = \frac{E_0}{t} = \frac{Sq_{r_0}}{t} = \frac{rq_{r_0}}{q_r}$$ [6]

Actual physical waste per period at production rate r per period plus exploitation (or depletion) per period at production rate r per period $= \frac{rq_{r_0}}{q_r}$.

[6] Stated in terms of actual physical waste :

$$\frac{Sq_{r_0} - Sq_r}{t} + \frac{Sq_r}{t} = \frac{Sq_{r_0}}{t} = \frac{rq_{r_0}}{q_r}.$$

This quotient can be broken down as shown above as follows :

$$\frac{rq_{r_0}}{q_r} = \frac{r(q_r+q_{r_w})}{q_r} = r+\frac{rq_{r_w}}{q_r}$$

This means if $q_r = q_{r0}$, then $q_{r_w} = 0$, hence,

$$\frac{rq_{r_0}}{q_r} = r = r_0,$$

or,

$$r+\frac{rq_{r_w}}{q_r} = r = r_0,$$

the reservoir is depleted at the optimum production rate and there is no physical waste.

The quotient $\frac{rq_{r_0}}{q_r}$ indicating current depletion and its associated actual physical waste (under the given production rate per period and the other conditions of the pool) is used in the cost function for crude oil production shown below.

(D) Depletion and Waste Costs

The costs for depletion and waste depend on the expected price for crude oil at the end of the lifetime of the pool under the given production rate.

Depletion and waste costs $= a_2 r q_{r_0}/q_r$,

where

$$a_2 = \frac{P_t}{(1+i)^t}$$ [7]

a_2 = Discounted price (present value of the expected price).

P_t = Expected price for crude oil at the end of the lifetime of the pool.

i = Interest rate expressed as a decimal fraction.

t = Lifetime of the pool, i.e. number of time units of exploitation under the given conditions.

Hence the costs for crude oil production can be expressed as the following function :

Crude oil costs of production =

$$(B_0+B_1d+B_2d^2)\frac{1}{t}+a_0+a_1r+a_2\frac{rq_{r_0}}{q_r}$$

[7] P_t is used because taking the oil now compromises only the ability to get oil at the end of the life of the pool, and not at any other time. Taking more or less oil now does not affect the possibility of getting out as much oil as it pays to get at any future time of high prices ; it merely brings nearer the time when production will stop entirely.

Drilling costs are here divided by the lifetime of the pool (t), where t itself depends on the production rate (r). In particular:

$$\frac{Sq_{r_0}}{t} = \frac{rq_{r_0}}{q_r}; \quad t = \frac{Sq_r}{r} = \frac{E_r}{r}$$

This means that depreciation charges are combined with the production rate.

(E) Empirical Investigation of the q_r Relationship

In August 1946 an analysis was published by R. C. Craze and S. E. Buckley[8] in which these authors investigated the data of seventy-six water drive reservoirs and twenty-seven gas drive reservoirs concerning the relationship between ultimate recovery and well spacing. Ultimate recovery was measured in " residual oil saturation expressed as percentages of total pore space," and the measure for well spacing was " acres per well." With reference to the ultimate recovery measure, a high residual oil saturation is equivalent to a low recovery efficiency and vice versa. In connection with these questions the following comments were made: " The shrinkage of the oil as gas liberated from solution affects the volume of the oil recovered in the stock tank. Fortunately, for water drive fields the per cent of the pore space in the reservoir left filled with oil at depletion is substantially independent of these volumetric effects and can thus be used as an index of recovery efficiency. The indicated ultimate recoveries of the water drive fields have, accordingly, been converted to residual oil saturations, expressed as percentages of the total pore space. For the fields considered, these residual oil saturations range between 17·9 and 60·9 per cent."[9]

The investigation mentioned was carried out with the technique of simple correlation analysis by successive elimination of the variables viscosity, pressure decline, and permeability. The result of the net-regression of residual oil saturation on well spacing is described in the following way: " there is no evidence in these data of a direct relationship between ultimate recovery and well spacing for the water drive sand fields studied."[10]

[8] R. C. Craze and S. E. Buckley: " A Factual Analysis of the Effect of Well Spacing on Oil Recovery," *The Oil and Gas Journal*, August 24, 1946, pp. 97—116.

[9] *Ibid.*, p. 106.

[10] *Ibid.*, p. 113.

The statistical data for seventy-six water drive reservoirs on which the Craze and Buckley investigation was based are used in this paper's multiple correlation least squares analysis. In doing this analysis it turned out to be useful to express the independent variables in terms of their respective logarithms.

The following relationship between the variables included was assumed :

$$X_1 = f(\log X_2, \log X_3, \log X_4, \log X_5)$$

Hence the linear multiple correlation equation appears to be :

$$X_1 = b_0 + b_2 \log X_2 + b_3 \log X_3 + b_4 \log X_4 + b_5 \log X_5$$

$X_1 = q_r$; efficiency rate.
X_2 = Well spacing in acres per well.
X_3 = Reservoir oil viscosity at initial conditions in *cp*.
X_4 = Initial reservoir pressure in *psig*.
X_5 = Permeability in *md*.

For the determination of our q_r values for the water drive reservoirs, some recalculations had to be performed : The ultimate recovery rates of the investigation mentioned are given as barrels per acre foot which, in our terms, is equivalent to E/af. The corresponding quantities " original total supply of crude oil in place " in barrels per acre foot (S/af) were not given but had to be computed on the basis of the information supplied. The computations were done according to :

Original total supply of crude oil in place =

$$S/af = \frac{C_0 p(1-w)}{V}.$$

C_0 = Number of barrels in one acre foot (7,758·372).
p = Porosity in per cent.
w = Interstitial water in per cent.
V = Oil volume factor (initial reservoir bbl. per stock tank bbl.).

Hence the efficiency rate is given by :

$$q_r = \frac{E/af}{S/af}$$

The values for q_r computed on this basis range between 0·2381 and 0·7820. Concerning the regression coefficients of our linear multiple correlation equation, the following signs could be expected :

(1) $b_2>0$. The density of the wells drilled on different reservoirs can be used as an indirect measure of different production rates at which these reservoirs are producing if the reservoirs in question operate each well with the same production rate per period within certain limits. This means a high number for well spacing (measured in acres/well), i.e. more acres per well or fewer wells in a given field is equivalent to a low production rate in that field and vice versa.[11] Hence q_r should be expected to increase as well spacing increases.

(2) $b_3<0$, i.e. the efficiency rate (q_r) decreases as viscosity increases.

(3) $b_4>0$. If the initial reservoir pressure increases, the q_r-values should be expected to increase.

(4) $b_5>0$. Increasing figures for permeability of the reservoirs should cause increasing q_r-values.

According to least squares multiple correlation technique, the following regression coefficients were obtained :

$$X_1 = 0{\cdot}346852+0{\cdot}103553 \log X_2-0{\cdot}158157 \log X_3-0{\cdot}118190 \log X_4+0{\cdot}157889 \log X_5$$

	Standard errors of estimate of regression coefficients	*Coefficients of partial determination*
$R^2 = 0{\cdot}508157$	$S_{12} = 0{\cdot}04113$	$r^2_{12{\cdot}345} = 0{\cdot}09699$
$\bar{R}^2 = 0{\cdot}474812$	$S_{13} = 0{\cdot}03551$	$r^2_{13{\cdot}245} = 0{\cdot}25166$
$S^2 = 0{\cdot}00900613$	$S_{14} = 0{\cdot}08080$	$r^2_{14{\cdot}235} = 0{\cdot}03499$
$S = 0{\cdot}09488$	$S_{15} = 0{\cdot}02780$	$r^2_{15{\cdot}234} = 0{\cdot}35355$
$N = 64$ [12]		

In evaluating these results, the following explanations might be helpful : The adjusted and unadjusted coefficients of multiple determination ($\bar{R}^2$, R^2) indicate the percentage to which the variance in the dependent variable (X_1) is explained by the independent variables ($\log X_2$, $\log X_3$, $\log X_4$, $\log X_5$). In our first study approximately 47·5 per cent of the variance in X_1 is explained by the variables included, whereas 52·5 per cent of this variance could not be explained.

[11]The term " production rate " used in this paper always refers to the amount of crude oil produced by all the wells of a certain reservoir.

[12]Both of the multiple correlation analyses are based on less than seventy-six observations. This is due to the fact that the data for the variables included in these analyses were partially incomplete.

The coefficients of partial determination ($r^2_{12\cdot345}$, $r^2_{13\cdot245}$, etc.) are measures for the individual importance of the independent variables and indicate the extent to which that part of the variation in the dependent variable which was not explained by the other variables can be explained by the addition of the new variable. The standard error of estimate of a regression coefficient is a measure of its significance. It indicates the range of variation of the coefficient. If this variation is large, it means that the significance or importance of the coefficient itself is low.

The term " significance level " refers to the probability that a sample will result in rejection of the hypothesis that the coefficient is zero (null hypothesis), calculated under the assumption that this hypothesis is true.

Concerning the interpretation of these results, first of all the attention should be called to two facts.

(1) The signs of the regression coefficients agree with the expected signs except for b_4. It should be noted, however, that the standard error of the coefficient $b_4(S_{14})$ comes rather close to the value of the coefficient itself, which means that its effect is practically zero. In other words, the pressure present in a reservoir *at the beginning* of its depletion does not affect significantly the efficiency rate. This variable could thus have been excluded from the analysis.

(2) In contrast to the investigation of Craze and Buckley, which used simple correlation analysis with successive elimination and which found no relationship between ultimate recovery and well spacing, the least squares multiple correlation technique performed on the basis of their data reveals a positive relationship between the ultimate recovery measure (q_r) and well spacing. This result confirms the hypothesis that the efficiency rate decreases as production increases. Concerning the reliability of the coefficient b_2, it can be noticed that it is more than two and one half times larger than its standard error, which means that it is significant around the 1 per cent significance level. The coefficients b_3 and b_5 are even more highly significant.

The difference between the two analyses might be mainly due to the fact that a simple correlation technique with successive elimination does not take account of the simultaneous interactions

between the respective variables. These simultaneous interactions between the independent variables are mathematically explained in a least squares multiple correlation analysis. The application of the successive elimination technique can give wrong results, particularly in cases where the partial correlation of certain variables is not high.

In order to corroborate the importance of well spacing, a second analysis was performed, in which pressure decline of the reservoirs (measured as percentages of initial reservoir pressure $= X_0$) was related to well spacing (X_2) and depletion to date (measured as percentages of ultimate production $= X_6$). Consequently the linear multiple correlation equation has the form :

$$\log X_0 = b_0 + b_2 \log X_2 + b_6 \log X_6$$

On the basis of the empirical data the following regression coefficients were obtained :

$$\log X_0 = 1{\cdot}328126 - 0{\cdot}765323 \log X_2 + 0{\cdot}504241 \log X_6$$

	Standard errors of estimate of regression coefficients	*Coefficients of partial determination*
$R^2 = 0{\cdot}415128$		
$\overline{R}^2 = 0{\cdot}396851$		
$S^2 = 0{\cdot}155207$	$S_{02} = 0{\cdot}1861$	$r^2_{02\cdot6} = 0{\cdot}2089$
$S = 0{\cdot}3939$	$S_{06} = 0{\cdot}1941$	$r^2_{06\cdot2} = 0{\cdot}0954$
$N = 67$		

This result again shows evidently the importance of the density of the wells drilled concerning influences on the current pressure situations of the reservoirs. The coefficient of $\log X_2$ in this analysis is more than four times its standard error, which makes it statistically highly significant.

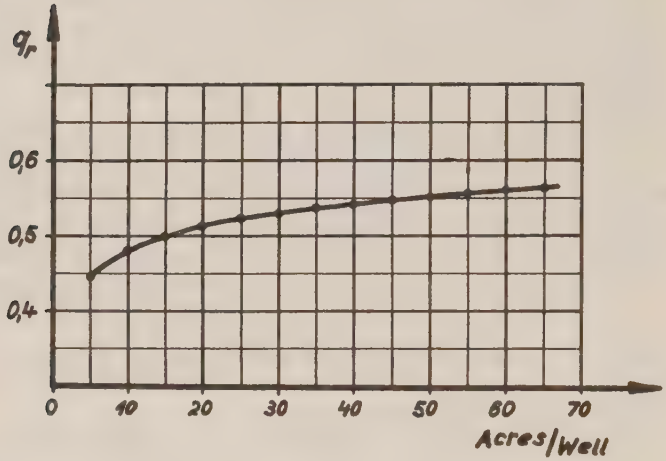
On the basis of the equation $X_1 = b_0 + b_2 \log X_2 + b_3 \log X_3 + b_4 \log X_4 + b_5 \log X_5$, the corresponding q_r-values were computed for well spacing ranging from 5–65 acres per well holding $\log X_3$, $\log X_4$, $\log X_5$ constant at their averages over the sixty-seven reservoirs studied :

$$\log \overline{X}_3 = 0{\cdot}1269$$
$$\log \overline{X}_4 = 3{\cdot}3054$$
$$\log \overline{X}_5 = 2{\cdot}7887$$

The results of these computations are shown in Graph 4.

GRAPH 4

RELATIONSHIP BETWEEN $q_r(X_1)$ AND $\log X_2$ FOR $(\log X_3, \log X_4, \log X_5) = \text{CONSTANT}$



The slope of the curve in Graph 4 agrees with the hypothesis set up in part C that $q_r = f(r)$ has to be downward sloped, since it is assumed that well spacing is an indirect measure of the production rate and the production rate is inversely proportional to well spacing, i.e. a high number for well spacing is equivalent to a low production rate and vice versa. Unfortunately, the statistical data published by Craze and Buckley contain no production rates for the reservoirs studied. Therefore, the relationship between q_r and r had to be measured indirectly by well spacing.

However, both of this study's multiple correlation analyses do indicate the existence of a negative relationship between the efficiency rate (q_r) and the production rate, even though production was measured indirectly by well spacing.

The results of this study can be improved by investigating the empirical relationship between well operating costs and the production rate. Furthermore, in the analysis on the relationship of the q_r values the variable X_2 (well spacing) should be replaced by the production rates at which the fields investigated were depleted.

This paper is the result of work done during my one year " Research Assistantship " at the University of Chicago, for which I express my deep gratitude to that University and its Department of Economics. I am furthermore indebted to *Professor Martin J. Bailey* for his personal help and technical advice.

GERHARD BIEDERMANN

University of North Dakota.

Measurement of Quality Changes in Consumer Durables

I. Introduction

This paper exhibits some of the tentativeness of a report on research in progress.* Nevertheless it develops notions of a theoretical character which, together with an empirical survey, may have broad interest. The central conclusions are three in number :

(1) if one distinguishes between consumer goods as sources of services and the service streams proper, intertemporal comparisons of costs of living become more feasible than has heretofore been suggested ;

(2) forms of industrial organization have much bearing on the appropriateness and effectiveness of procedures that might be undertaken to measure quality changes of sources of services. Specifically, we are concerned with the perfection of second-hand markets, concentration ratios, and the degree to which there are important indivisibilities in the production process. Supply considerations cannot be neglected in cost-of-living studies. Ideally-organized second-hand markets make possible precisely-calculated " improvement factors," measures of quality change ;

(3) empirical investigation of the retail market for household refrigeration in the United States indicates that the failure of the U.S. Bureau of Labor Statistics to account for changes in quality (including size) may have imparted a downward bias to indexes computed up to 1956.

II. Sources and Services

One of the perplexing difficulties of our topic is rooted in the fact that a given machine is usually the source of a number of services. An automobile is a source of transportation services, satisfaction arising from invidious comparisons, and of aesthetic sensation ; it serves as a money substitute to the extent that it is a store of value. It is rare that the service yield of a consumer durable can be described in terms of a single dimension ; complex

*I am greatly indebted to my assistant, Timothy G. Dalton, Jr.

numbers must be employed. The fact that service yield is multi-dimensional in character is immediately suggestive of a key difficulty : a vector a is greater than another vector b only if each element of a exceeds the corresponding element of b. In longer-run comparisons between machines of different data, it is likely that standard machines will exhibit " mixed " changes. Some elements of the original service vector that were non-zero will become zero or more negative and vice versa. It is obvious that the longer is the interval separating the points in time for which comparisons are desired the less likely is it that we can find substantial numbers of models that are identical. And this, when all is said and done, is the quality problem.

Over the period 1935—60, for example, we observe massive changes in newly produced and initial stocks of sources of services. In 1960 refrigerators are larger, have substantial freezer space, and frequently feature automatic defrosting. Automobiles are much different in appearance and in many performance characteristics. Food products sold at retail often are much nearer to the final-product stage. Television sets existed in large numbers in 1960 whereas TV was commercially unavailable in 1935. It is clear that vectors representing 1960 products markedly differ from those representing 1935 products. The quality of many 1960 products is found to be profoundly different from that of 1935 counterparts.

But is the world so different a place? If we reduce the problem to one of comparison of service streams, might we not find that the 1960 and 1935 worlds are much more different in terms of techniques employed to yield service streams than in terms of the *nature* of the streams? I think the answer is " yes." Take the case of television. Certainly there would not appear to be a qualitative change more striking than the post-war development of TV. But what *services* does a television set render? Generally speaking, entertainment services of various sorts : viewing of athletic events, vaudeville shows, plays, movies, etc. Of course, television has caused profound changes in the life of modern societies, but these have been changes in the way services are rendered ; the constituent services are substantially the same, although the appropriate weights might differ.

It follows that at least crude imputations of the carrying and running costs of TV could be made to the various service flows (it

being unnecessary to point out how pay-TV would simplify these tasks). Perfectly reasonable comparisons can be made between observing an athletic contest in 1935 from a " cheap " seat and watching a game on TV in 1960, due attention being paid to the cost of transportation to and from the game in 1935. Problems of weighting arise, but it can be argued that TV viewing of baseball in the United States, for example, was immensely more important in 1960 than were bleacher admissions, that the latter can be ignored on the same ground that 1960 indexes of the cost of household refrigeration ignore purchases of special-order ice boxes. It might be argued that observing an athletic event via TV is *not the same* as observing it in a cheap seat. True enough, but Manolete is not Ordonez ; Tilden is not Hoad. Time–series comparisons never achieve metaphysical likeness. Cost-of-living indexes can never be more than heuristically appealing.

The case of food is, of course, something of a digression in a paper concerned with durables, but is interesting both because it is easy to conceptualize and because it suggests an important caveat : we must beware of pestling the concept of service yield so fine that the resulting measures are without economic significance. Thus, we might reduce food consumption to intake of calories, protein, etc., and arrive at a solution of this self-imposed diet problem by determining how much it would cost in 1960 to acquire 1935 nutrient values in the cheapest way. The answer would have little to do with economic behavior or the solution of economic problems, since there are many other qualities of food important for economic decision making. On the other hand, it is not terribly difficult to compare the cost of serving a given quantity of spinach in 1935 with that in 1960, even though the basic ingredient was purchased only in its natural state in 1935 and in ready-to-cook form as well as in its natural state in 1960. We must, of course, impute a value to the labor saved in the newer preparation. What of spinach sold in 1960 in its natural state ? It might be found that the base price for spinach in 1960 is not equal to the difference between the price of the frozen product (for example) and the imputed costs of household labor devoted to processing. Such a discrepancy could develop from differences in habits and tastes among consumers. Under present procedures of the United States Bureau of Labor Statistics, it

is unlikely that the requisite weighting scheme would be employed; a 1935—60 comparison would be confined to foods pretty much in the same category of " naturalness." While it is true that only such comparisons can achieve metaphysical nicety, it is also true that such nicety sacrifices relevance and leaves the quality problem unsolved.[1]

III. The Impact of Forms of Industrial Organization

It is not uncommon, certainly in long-run comparisons, to find that some models within a given product category have been withdrawn by manufacturers and that others have been introduced. The implications of forms of industrial organization for the kinds of inferences that can be drawn from such phenomena have been neglected.

Take first a hypothetical case of the following type: (1) in period $(t-1)$ only 8-cubic-foot refrigerators are sold and in period (t) only 10-cubic-foot machines (ignoring carry-over and the second-hand market); or (2) let us assume that a given source yields a service measurable in such terms as " photon per minute " but that the use of the source involves a certain amount of risk and that there comes to be developed a device that, when incorporated into the source, meaningfully reduces the risk. Assume further that the device was unavailable in $(t-1)$ but that all machines sold in (t) are equipped with it.

There are some circumstances which, when associated with competitive markets, enable us to make unambiguous assertions about the direction of price change. Thus, if the (t) machines sold at the same price as did the $(t-1)$ machines in period $(t-1)$, and if the new machines were no heavier and occupied no more outside space, price relatives based on almost any sensible criteria would be expected to be less than 1·0. Take the more interesting assumption that $P(t)$ exceeds $P(t-1)$ by the difference in cost occasioned by the change, assuming no change in factor prices.

[1] The phrase, " substantially-identical service streams," shields a host of difficulties. The easiest case for us is one in which the service stream is physically measurable in some " photon per minute " sense. The hardest is one such as furniture in which the service stream is not physically measurable and in which tastes obviously are important. Another extremely difficult case would be oil paintings. In these instances to classify streams as " substantially identical " is to slash wildly with Occam's razor and rip the economic content out of the problem.

Under certain circumstances, one again can infer that the price relative should fall. Particularly, if machines of any subtype were produced at constant average cost throughout and if the industry were purely competitive (not pausing to establish another condition or two so as to determine the size of the firm). In this case anyone who wished to purchase a $(t-1)$ type machine could do so at a price equal to $P(t-1)$. Nobody chose to do so. It follows that nobody in the market valued the improvement at less than the change in price.

But the inference that the new feature is more highly valued by all consumers than the cost differential cannot be drawn, even under the above cost conditions, if the market is monopolized or featured by oligopoly. A simple little model can illustrate this. Assume that the source category X is monopolized in accordance with the following joint demand and cost functions :

(1) $q_1 = f_1(p_1, p_2)$
(2) $q_2 = f_2(p_1, p_2)$
(3) $g(q_1, q_2) = C$

Ignore the facets of a more sophisticated approach : recognition of the effects of sales of sources on future demand ; determination of optimal durability ; promotional activity, etc. It is apparent that X_1 and X_2, the basic and advanced models, are substitutional. An increase in p_1 will increase the maximum price consistent with a given rate of sale of X_2, q_2. It might turn out that the marginal rate of gain from increasing p_1,

$$q_1+p_1(\partial q_1/\partial p_1)+p_2(\partial q_2/\partial p_1)-(\partial C/\partial q_1)(\partial q_1/\partial p_1)-(\partial C/\partial q_2)(\partial q_2/\partial p_1),$$

is positive for all p_1 at which q_1 is positive, given that p_2 is always chosen to maximize profits, given p_1. Profit maximization proves inconsistent with the production of X_1. And, of course, our conclusion is strengthened when there are meaningful indivisibilities of production, since it is then possible that, even if first and second-order conditions for profit maximization are met with positive q_1 at a moment in time, the firm might be better off in the long run by abandoning the production of X_1 as resources specialized to its production wear out.

Within the context of the monopoloid-oligopoloid model, it can no longer be assumed that there are not consumers who would have preferred the former state of affairs. The " losers " would

include consumers who withdrew from the market as well as others who accept the new arrangement " at the cost of their consumer surpluses."

Another set of problems hinging on market organization pertains to the existence and perfection of second-hand markets. It is impossible to discuss problems of measuring price changes of durable goods without introducing second-hand markets. These markets offer opportunities for the analyst but also pose new difficulties. For example, it must be confessed that all of the proposed applications of data from second-hand markets involve a highly questionable assumption about the calculability of depreciation rates, namely that there is an identifiable depreciation rate. Among possible applications of these data, three are considered here :

(1) continuation of pricing of items discontinued in production for cost-of-living purposes ;

(2) adjustment of such indexes where production of " old " items persists but where cost conditions induce " distortions " ;

(3) use of the improvement-factor technique to achieve measures of improvement in quality.

(1) *Use of items discontinued in production in cost-of-living indexes*

This is the most obvious application. We see at once that the existence of data from second-hand markets permits us to price machines at (t) identical with those priced at ($t-1$) except for age, and that the identifiability of a physical depreciation rate would make it easy to compare otherwise equivalent sources.

(2) *Adjustment for distortions induced by cost conditions*

Attention should be paid to what can be a seriously misleading constellation of price quotations in a period of model change. Assume, for example, that the safety feature of the t-model refrigerators just discussed has resulted in an increase in price over that of standard machines of $10, but that there are some consumers who value the change at less than $10 *and* who dislike the idea of using a second-hand machine. It follows that a certain number of old-type machines will continue to be produced and sold in (t) at a price of, say, $300—assumed to be the price of such machines in ($t-1$)—in contrast with the $310 price for new

machines. Now assume that many other customers have no prejudice against used equipment ; these have a lower offer price for used machines only because of physical depreciation and are unwilling to pay a premium for newness, *cet. par.*, and have a strong preference for the safety feature. It follows that these persons will plan to unload their old-type machines and switch to the new-type machines if the price of the former is quoted in the second-hand market at $\$(1-k)$ (300). It is altogether possible that the \$300/\$310 price pattern (in the " new " market) will be accompanied by circumstances finding holdings of old-type sources planned for the end of period (t) to be less than $(1-k)$ (S_{t-1}), where S_{t-1} gives the number of sources, expressed in stored service units, available at the end of $(t-1)$, the beginning of (t). It follows that the price of $(t-1)$ sources will fall below $\$(1-k)$ (300).

We find then that a substantial disparity can come to prevail between the cost of production of previously-standard sources and their price in second-hand markets, that the price of second-hand sources can fall far below the supply price of new sources of that design, much farther than mere depreciation considerations would suggest. Cost conditions in the new-source market prevent this change in relative price from being reflected in supply prices of new output. The failure to consider second-hand price quotations might give inordinate weight to the market transactions of those with strong prejudices against second-hand machines.

(3) *The improvement factor technique*

It is open to us to achieve quantitative measures of quality change if we are willing to rely on a sort of collective judgment through a simple technique which we christen the calculation of *improvement factors.*

Examine prices paid for new-type sources (first produced in t) in second-hand markets, noting that current models find their way into these markets early in the model year (although we must exercise caution to the extent that " lemons " are likely to be the first to so enter). If the " market," whatever that means, " feels " that the producer's claims are without merit—to take the monopoly case—we would expect that the price of new-used models would exceed that of models of, say, period $(t-1)$ only by a factor

given by the depreciation rate, k, the factor being $1/(1-k)$. Once this factor is determined, if we wish to confine ourselves to new machines, we can compute the improvement factor according to the formula

$$IP = (P_2/P_1)\ (1-k),$$

where P_2 is the price of a new-used source and P_1 that of a $(t-1)$-model, both observations being made during (t). To the extent that IP exceeds 1·0, it indicates an " increase " in quality.

In fact, the IP technique would appear to represent a way out of the difficulties posed by the set of problems flowing from changing service-yield vectors. It permits us to concentrate on readily-observable things such as machines rather than on what might be incorporeal figments, service-yield streams. There are difficulties :

(1) the technique relies on what might be fragmentary second-hand market quotations ;

(2) its application depends upon k being observable. Where service-yield can be measured in " photons per minute," k might be easy to compute, but in more realistic cases we must estimate k from second-hand market data which might make difficult identification of k as against past improvement factors. A high value for k and continuous improvement of product can easily be confused unless there has been a meaningful number of periods in which standard outputs remained substantially unchanged in service-yield characteristics ;

(3) the method gives great weight to the judgment of what might be sucker types (although an alternative speculation suggests it to weight heavily resistance to change among buyers). A caustic critic might maintain that the observations were generated by a process like the following : disillusioned purchasers threw their machines into the market and were pleased to discover that there were some consumers hopelessly victimized by the blandishments of Madison Avenue. The great majority of stockholders within the source category failed to succumb to these blandishments in the first place—they did not purchase this year's models—and certainly were

not willing to purchase new-used models of vintage (t) except on a strongly discounted basis, but the original victims were able to pass on their misery in a thin market.

An application of the improvement-factor technique to a problem substantially untouchable by other techniques such as hedonic indexes might be helpful.[2] Consider durable goods yielding service streams the elements of which are not quantifiable but which are unlike such goods as paintings in that changes in taste are not critical ; it may be assumed that response to a given set of items within the product category depends upon traditional economic variables subject only to trends affecting the *entire* category. Let us concentrate on sporting equipment, specifically golf clubs.

Over the years, important improvements have taken place in the industry, it being almost unanimously agreed, and substantiated by statistical evidence, that a golfer of a given level of skill, playing under otherwise identical circumstances, will be able to lower his score by using modern clubs. On the other hand, the bases for improvement—lamination of wooden club heads, tempered-steel shafts with greater whippiness, matched clubs, etc.—are not quantifiable.

Considerable care must be exercised in employing the improvement-factor technique. But if we know the " normal " depreciation rate (on declining balances) and wish to compute a price relative for standard sets of golfing equipment for the years $(t-5)$ and (t), we take the product

$$\frac{P_a(t)}{P_b(t)} \cdot (1-k)^5$$

so as to determine the improvement factor (where the improved model is a and where it is assumed that model b ceased to be

[2]Compare A. T. Court, " Hedonic Price Indexes with Automotive Examples," *The Dynamics of Automobile Demand* (Detroit : 1939), pp. 99ff. Also R. Stone, *Quantity and Price Indexes in National Accounts* (OEEC : Paris, 1956), Chapter IV. An important further contribution is made by Zvi Griliches in a staff paper for the Price Statistics Review Committee to be published by the Joint Economic Committee of Congress (Washington : 1961). Cf. also a forthcoming paper on the theory and practice of hedonic indexing by Zvi Griliches and Irma Adelman to be published in a forthcoming issue of the *Journal of the American Statistical Association.*

produced after $(t-5)$ while model a first came into production subsequent to $(t-5)$). If this ratio were 1·5 and if the ratio

$$\frac{P_a(t)}{P_b(t-5)}$$

were also 1·5, we would define the price relative as 1·0, the improvement factor exactly compensating for the increase in nominal price. However, at least two caveats must be observed :

(1) for a given distribution of tastes among the population, as the older equipment wears out, the price of clubs of a given vintage will increase relative to the price of new clubs, *cet. par.*, to the extent that certain peculiarities of older equipment are not found among newer clubs (wooden shafts for example). This might be offset by the effects of increased familiarity with new equipment and by a higher depreciation rate for human as against non-human agents, but, nevertheless, as time passes, we are in danger of pricing goods valued as antiques.

(2) sometimes the novelty of the new equipment might result in the calculated improvement factor being unduly high or low depending on the enthusiasm of the initial response. Moreover, if the market is not perfectly competitive (and cases dealing with innovation can hardly be fitted into such a framework) we might find that the new item was introduced in a very small quantity with its price correspondingly high, suggesting that we should wait a bit before using data from second-hand markets. This consideration can be important even in substantially competitive markets when scale economies are significant.

We have not hesitated to attach rather severe *caveats* to our revelation of the *IP* technique. Nevertheless, it is considered to offer the first real opportunity to make creative use of data from second-hand markets in an effort to maintain continuity for longer-run comparisons and to achieve quantitative measures of quality change.

IV. A Paradoxical Empirical Application

It is often taken for granted that " naive " pricing techniques tending to ignore quality changes have the effect of introducing an

upward bias into price statistics. It is shown in this appended section that such is not necessarily the case. Quite the contrary.

The procedures followed by the U.S.B.L.S. in developing refrigerator components for the Consumer Price Index have been much simpler than those speculated about above. At any given time, we would find the Bureau pricing a single model, considered representative of the " average " expenditure of appropriate city residents. From time to time the specifications are changed in response to changes in model mix. At such points a linkage is made.

The " standard " item or the subcategory for which it stands is often a surprisingly small part of total expenditure. Thus, over the 1948—51 period, the Bureau priced an 8-cubic-foot refrigerator. At the outset (1948) only 29 per cent of refrigerators sold exceeded 8 cubic feet in capacity, but, by 1951, 47 per cent of machines sold did so, and, of course, an even greater proportion of *expenditure* on refrigerators was accounted for by the larger, more expensive machines. At the same time, freezers came to account for an ever-larger volume of sales and were not priced in the Bureau's refrigerator price index. Thus, in 1951, 1,050,000 freezers, valued at $310,300,000 were sold in the U.S. while the total value of freezers and refrigerators sold that year was $1,150,000,000.[3] In May 1952 the specification was changed so that machines of 8·0—9·7 cubic feet were priced, this size specification holding until January 1955 when it was increased to an 8·8—9·7 cubic foot range. In February 1956 the specification was again changed so that the size range was 10—11·5 cubic feet. That specified for April 1960 was 10—12·2 cubic feet. In 1952, 45 per cent of refrigerators sold were 8 cubic feet or less and 21 per

[3]The refrigerator–freezer price–output data are taken from my paper " The Demand for Household Refrigeration in the United States," Chapter III (pp. 99—145) of A. C. Harberger (ed.), *The Demand for Durable Goods* (Chicago : 1960). See especially pp. 102—108, 13—135.

Mr. Arnold Chase of the BLS kindly supplied details of BLS specifications for electric refrigerators.

In an effort to measure realized instead of more-or-less fictional suggested list prices, I relied on Sears-Roebuck mail order quotations and on average sales realizations of that giant concern, including sales of its many retail stores.

cent exceeded 10 cubic feet in capacity. Freezer sales in 1952 accounted for $367·4 million with combined refrigerator and freezer sales of $1·16 billion. In 1954, 33 per cent of refrigerators sold exceeded 10 cubic feet in capacity and freezer sales accounted for $335·5 million out of total sales of $1·29 billion.

We become concerned with the extent to which we are justified in assuming that price movements of items selected for pricing reflect movements of items that are related but excluded. If it could be demonstrated that the simple procedure attributed to the Bureau leads to the same results as do more elaborate and costly procedures, the Bureau procedures could be supported on grounds of economy if nothing else.

I have fully described elsewhere the procedures followed in construction of an alternative price index for household refrigeration in the United States over the 1931—55 period.[4] These procedures were intended to reflect actual prices paid to a greater extent than would other procedures and were formulated to catch the effects of quality changes to the extent that

(1) Laspeyres and Paasche indexes were computed with the weights for different size categories shifting annually according to market behavior. The index was constructed through annual chain links ;

(2) freezers were included ;

(3) within each size category, the models compared were in most instances selected so as to be as much alike as possible between any two periods.

My indexing technique reflected rather more consciousness of the quality problem than did the Bureau procedure.

The resulting price indexes, together with a third index (column 3) applying Sears catalog quotations to machines meeting official specifications, follow :

[4]M. L. Burstein, " The Demand for Household Refrigeration in the United States," *loc. cit.*, note 3 *supra*.

TABLE 1

Year	(1) BLS Index	(2) MLB Index, P II	(3) BLS Spec., Sears Prices
1948	100·0	100·0	100·0
1949	96·9	89·1	80·1
1950	94·8	87·2	80·1
1951	99·8	88·1	81·8
1952	86·0	87·6	83·3
1953	92·4	81·1	68·9
1954	86·1	75·9	58·0
1955	80·9	71·9	56·2

Column 3 tells us what the BLS figures could be expected to be if its price-reporting procedures better reflected actual prices paid. In fact it reduces to an index ignoring quality changes but reflecting accurate reporting of prices of models chosen for inclusion. Comparison of the figures in columns 2 and 3 shows that the latter index fell substantially more than the former over the 1948—55 period ; the index ignoring quality changes fell more than the index adjusted for such changes with reporting procedures held constant.

Why should procedures ignoring quality change impart a *downward* bias to a price index ? A plausible explanation is that prices of the generally obsolescent models chosen for pricing by the Bureau fell relative to the prices of refrigerators and freezers as a group. The " BLS models " lost attractiveness to buyers faced with opportunities to purchase improved machines. This result contains the clear implication that mark-ups fell for the " BLS models " a result inconsistent with pure competition in the industry and once again causing us to stress the importance of forms of industrial organization for price-indexing problems.

M. L. Burstein

Northwestern University.

Notes on Cost and Capacity

In this paper I shall be concerned with relations between costs and capacity. Although it is potentially an important part of the theory of diminishing returns, there has been very little discussion of the concept of capacity itself. It might almost be said that economic theory has developed without a concept of capacity at all.

Diminishing returns—when?

That economic theory has operated without a definite concept of capacity is due to the nature of price theory, and to the rarity of periods of sustained full employment. Where, as in price theory, output is determined by the relation between costs and sales, and supply is elastic because there is plenty of slack in the system, there is little call to examine the physical characteristics of plant. Moreover, interest in the effect of reaching the limits of physical capacity has been effectively damped by the way economists have interpreted the *law of diminishing returns.*

According to this law, an increase in some inputs relative to other comparatively fixed inputs will cause output to increase; but after a point the extra output resulting from the same additions of input will become less and less.

Stigler refers to "the extreme generality of the law" [1] and uses it to justify the common practice of drawing marginal cost curves which rise from low levels of output. (*See* Diagram 1). Yet, as can be seen from the formulation of the law the critical point at which diminishing returns sets in is left undefined. Statements such as "at first the countertendency of increasing mass-production efficiencies may outweigh the tendency towards diminishing returns but ultimately diminishing returns will win out" [2] clarify the law in no way. There is nothing in the law to tell us whether the critical point is 30 or 90% of the physical capacity of plant. That is a question of fact that can only be established by empirical studies.

[1] *The Theory of Price.* Revised Edition, p. 112.

[2] Samuelson, *Economics,* 3rd Edition, p. 23.

Perhaps concentration on the theory of competitive equilibrium made it orthodox to assume that marginal costs per unit rise from low levels of output relative to capacity. In the perfectly competitive model, marginal costs must rise to establish equilibrium. Stigler has argued that if the marginal costs were constant over a wide range of outputs and then rose steeply for each firm, the output of the competitive industry would vary in the short run chiefly through variations in the number of plants in operation and hardly at all through variations in the rate of output of plants that stay in operation. But this, he argues, is the opposite of the facts.[1] Yet his deduction is valid only if there are significant interfirm differences in marginal costs and if there are no special markets due to such things as product differentiation and transport costs. Since such special markets are common Stigler's indirect proof of the generality of the law is far from general.

The relevance of the law is taken for granted in surprising situations. Even Keynes took it for granted that recovery from depression would involve a diminished real wage because there would be "decreasing return from applying more labour to a given capital equipment." [2] Occasionally however, Keynes points to (and at times inconsequentially relies on) factors that do not fit the classical law of diminishing returns. He refers to situations where there is "no surplus of specialized or practised labour and the use of less suitable labour involves a higher labour cost per unit of output" and he remarks that "this is the main reason why the supply price of output rises with increasing demand even when there is still a surplus of equipment identical in type with the equipment in use . . . —an outstanding case of rising supply price with increasing output not due to internal diseconomies." [3]

In his chapters on "The Theory of Prices," he mentions equipment that is non-homogeneous and implies that as output is increased, less efficient equipment will be brought into use, so increasing marginal prime costs. "Hence, in general, supply price will increase as output from a given equipment is increased.

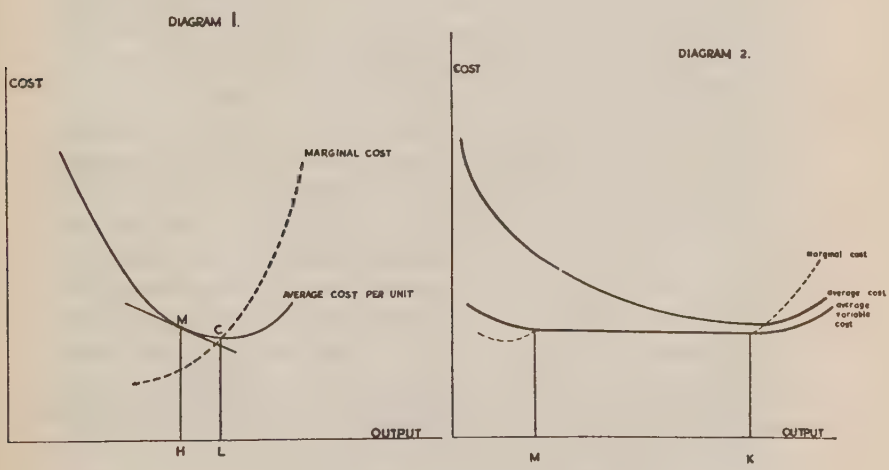
[1] *The Theory of Price*, p. 167.
[2] *The General Theory*, p. 289.
[3] *The General Theory*, p. 42.

Thus, increasing output will be associated with rising prices, apart from any change in the wage unit." [1]

The excessive generality of this conclusion is obvious. If there is no shortage of specialized or practised labour and if there is more equipment identical in type with that in use, increasing output will not necessarily be associated with rising cost.

Where it is not possible to increase output at constant cost, the extent to which cost will rise is a matter of interest. This is particularly so for a cost theory of the price level. In times of full employment, the concept of capacity takes on a new significance, which doubtless explains why recently economists have become much more interested in estimating the extent of spare capacity and in putting numbers to the degree of spare capacity required (or thought to be required) for growth without inflation.

There is a very specialized concept of capacity used in welfare economies. In the perfectly-competitive model of economic adjustment, the price of the product is equal to the lowest level of average cost in the marginal firm. Thus, in diagram 1, the equilibrium output of the marginal firms will be *OL* at price and average cost *CL*. This is said to be an



[1] *The General Theory*, pp. 299-301.

optimum situation. Where there is product differentiation but free entry, however, output will be only *OH* and price and average cost will be *MH*.

In Chamberlin's model, *HL* measures *excess capacity*.[1] By implication capacity is equal to *OL*, where average cost per unit is at its minimum. That not much can be done with this concept of capacity is clear from the fact that firms in monopolistic or semi-monopolistic conditions, may, in fact, work at or above Chamberlin's *capacity*. If the downsloping demand curve in diagram 1 is shifted to the right, excess capacity will be reduced and may become negative.

Ways of increasing output.

Generally, relations between cost and output will be less simple than those implied in diagrams 1 and 2. It is quite possible to have several short period costs for particular values of output, without even allowing for changes in factor prices.

Keynes pointed out that increased demand for output will affect cost through expectations. A rise in output that is expected to continue may raise marginal cost where a rise in demand, that is expected to be temporary, may not. This, as pointed out by Keynes, will happen because of differing effects of expectations about future events on marginal user cost. "Marginal user cost may increase sharply when employment begins to improve, if . . . the increasing effective demand brings a rapid change in the prevailing expectation as to the date when replacement of equipment will be necessary.[2]

Costs may vary without change in expectations. This is obvious when we consider the ways in which output may be increased. Suppose, first, that a firm is able to recruit unemployed labourers capable of working as effectively as those in employment, that it operates plant that is completely divisible, and that it has spare capacity. If the prices of all the variable factors are unchanged, output can be increased $x\%$ at constant marginal cost. If, however, the firm is only

[1] "Whenever price competition fails to function . . . the result is not merely higher prices, but also excess capacity as a permanent and normal characteristic of the equilibrium adjustment." Chamberlin, *Theory of Monopolistic Competition*, 1939, p. 171.

[2] *The General Theory*, p. 302.

able to recruit less effective workers, though otherwise conditions are the same as in the first case, marginal cost will rise because of the higher labour cost per unit output. If, finally, there are no effective workers available, increasing output by $x\%$ will entail overtime working at higher rates of pay. Here also, marginal cost will rise because of the higher labour cost per unit output.

An increase in output may also be costly because satisfactory materials or components are not available when required. If, for example, following a rise in demand, car manufacturers or builders responded more quickly than their suppliers, interruptions to the planned flow of work would raise real costs. The same situation would arise where growth in demand elsewhere led to demand in excess of the capacity of the firms supplying materials and components.

These variations in cost are determined by factors external to the firm. Let us now consider variations of a more constant and predictable kind. Consider first the problem of working overtime. Suppose that a firm produces a standard product and that its machines require constant inputs of labour and materials per hour. Such a firm will vary output by varying hours of work. Let diagram 2 represent the cost condition of such a firm, with K the output from a standard working week. For output less than K, workers would be put on short time. Output greater than K would require overtime which, with labour paid at overtime rates, would raise marginal costs. So long as the normal work is shorter than that required for labour efficiency it will, in given conditions of factor supply, require overtime rates to push up marginal cost. Overtime rates rise for weekend working and beyond a certain level fatigue causes labour per unit output to rise. It follows that beyond certain levels of output, the introduction of shift working would lower both average and marginal costs of production.

An example from cotton spinning will give a better idea of these complicated relations between cost and output. A cotton spinning firm could work overtime, or a short evening shift of 2–4 hours at normal wage rates, or double-day shifts in which "the hours in each shift are to be not less than 37½ and

TABLE I

COSTS IN SPINNING[1]

£ per week

	One Shift[2]			Short Evening Shift[3]		Double Shift[4]			
Hours ...	22½	45	50	47	49	75	79	82½	90
Heating ...	50	50	50	50	50	50	50	50	50
Power ...	213	325	350	335	345	475	495	513	550
REPAIRS									
Building ...	30	30	30	30	30	30	30	30	30
Machines...	95	140	150	144	148	200	208	215	230
Salaries ...	200	200	200	200	200	250	250	250	250
Wages ...	450	900	1,050	940	980	1,800	1,920	2,025	2,325
CAPITAL									
Working...	112	225	250	235	245	375	395	412	450
Fixed ...	2,960	2,960	2,960	2,960	2,960	2,960	2,960	2,960	2,960
Total Cost	4,110	4,830	5,041	4,894	4,958	6,140	6,308	6,458	6,845
Average Cost[5] ...	182	107	101	104	101	82	80	78	76
Marginal cost per unit ...	32	32	42	32	32	44	42	42	53

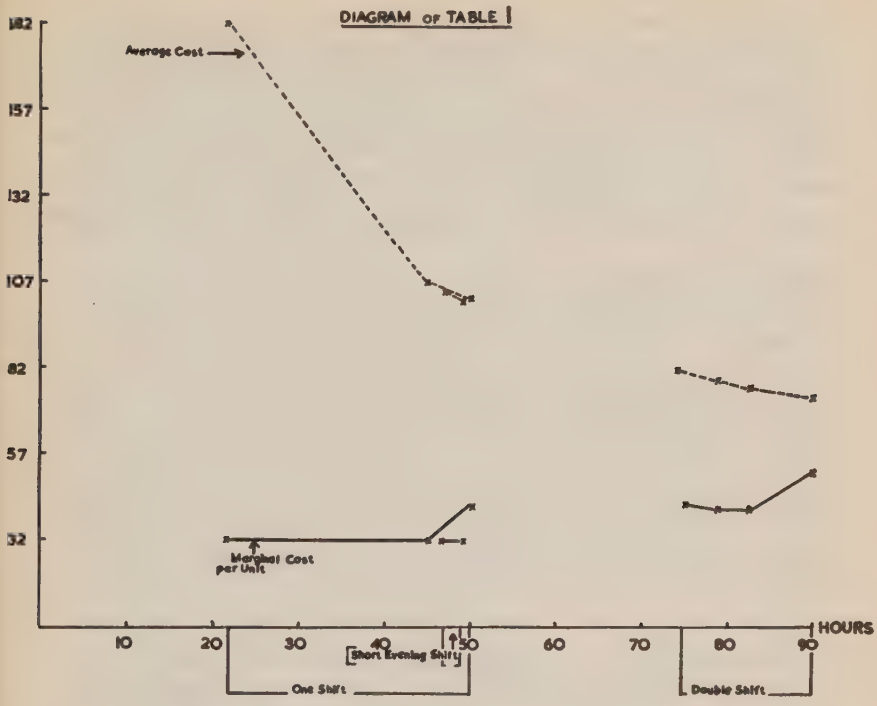
[1] These costs are developed from the data given in Robson. The important assumptions are that labour and materials per unit output are constant, and that it is appropriate to write off machine over ten years because it will be obsolete then. This makes depreciation independent of the rate of use.

[2] Wages at time and one half for hours in excess of 45.

[3] A short evening shift of 2 or 4 hours at normal wage rates.

[4] On double shift, hours are reduced from 45 to 37½ without reductions of pay. Saturday work is paid time and one half ; Sunday double.

[5] As labour per unit output is assumed constant, Total Costs are divided by hours to give an indication of the movement of average and marginal costs.



not more than 38¾ spread over 5 days with payment as for 42½ hours." [1]

In *The Cotton Industry in Britain* (1957) Dr. Robson gives figures for the weekly cost of running a spinning mill of 50,000 ring spindles at physical capacity and at half-capacity. In Table I, I have extended his figures on the basis of his assumptions about labour and materials per unit output to show the effect on cost of working beyond the capacity of 45 hours (which were the normal hours for one-shift systems when he wrote). By assuming that labour per unit output is constant, I have produced an index of average cost and marginal cost per unit. If labour per unit output is in fact greater for overtime and double-shift than for single-shift working, the rise shown in marginal cost for overtime and double-shift working will be understated.

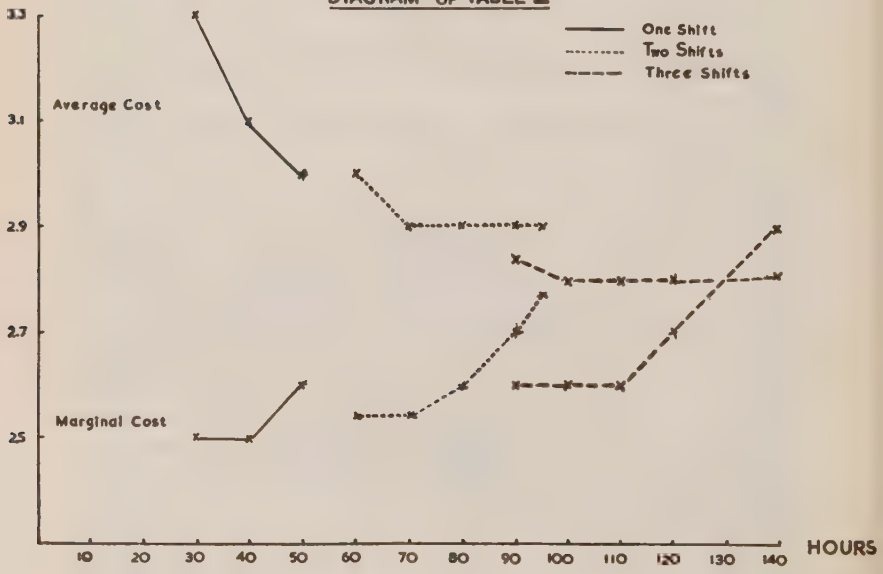
The movement of average and marginal cost per unit is plotted in the diagram to Table I. This makes clear the discontinous nature of the cost functions where there are possibilities of overtime and shift working.

[1] *Time Rates of Wages and Hours of Work, 1st April,* 1961. H.M.S.O.

TABLE II

	One Shift			Two Shifts				Three Shifts				
Hours ...	30	40	50	60	70	80	90	90	100	110	120	140
Labour ...	6	8	11	14·5	17	20	23·5	22·5	25	27·5	31	40
Materials...	60	80	100	120	140	160	180	180	200	220	240	280
Variable Overhead	9	12	15	18	21	24	27	27	30	33	36	42
Overhead	25	25	25	25	25	25	25	25	25	25	25	25
Total Cost ...	100	125	151	178	203	229	256	254	280	306	332	387
Average Cost ...	3·3	3·1	3·0	3·0	2·9	2·9	2·9	2·8	2·8	2·8	2·8	2·8
Marginal Cost ...	2·5	2·5	2·6	2·55	2·55	2·6	2·7	2·6	2·6	2·6	2·7	2·9

DIAGRAM OF TABLE 2



In Table II, I have used some rounded accounting figures provided by a firm in the motor industry. Labour and materials per unit output were shown as constant for a two-shift system, which was said to "reflect the theoretical behaviour of such variables, though in practice it was found that variable costs per unit rose as output fell." I have extended these figures to a three-shift situation though in fact the firm only uses three shifts in one department.

The movement in labour cost per unit output is due to changes in rates of pay in accordance with the rates agreed by the Engineering Employers' Federation and The Confederation of Shipbuilding and Engineering Unions. Overtime rates apply to hours in excess of 42 for one shift, 75 for double-day and 112½ for three shifts. On double-day or three-shift systems, the weekly hours of work are 37½, payment being made for 42 hours on each shift, but for workers employed on the third shift, there is a special bonus of 2/42.

As with spinning, the assumption of constant labour and materials per unit output has been used to derive an index of average cost from hours and total costs. It follows from the assumptions that movements in marginal cost are determined solely by wage premiums. In this case, marginal cost does not exceed average cost until there is overtime on a three-shift system. It is clear, however, that a relatively small fall in labour efficiency during the third shift would push marginal cost above average cost. Except in continuous process industries, efficiency on third shift is usually low and it could prove impossible to get workers on a special bonus of only 2/42. Third-shift working would, therefore, lead to a greater rise in marginal cost than shown in Table II.

So far there has been no mention of the effects of the *product mix*. Cost varies not only with the volume of production and the prices, qualities, and deliveries of inputs, but also with the size of production runs, the variety of output and the level of selling costs. To treat the cost function as a relation between cost and volume of production only, is pointless outside conditions of perfect competition.

Because the product mix may have a marked effect on cost, a given value of output may be associated with several

different cost levels. Where the market situation is not entirely favourable, a firm may accept orders which it is not adapted to produce efficiently or on which profit margins are low, or it may increase variety or promise quick delivery of all its products, with a consequent increase in costs due to a rise in the proportion of make-ready time. By contrast, in times of excess demand for its products, the firm can choose the most profitable lines, reduce variety, and/or set delivery times that allow it to arrange long production runs. There must be many firms where the correct answers to questions about costs and output would specify the ranges within which costs would move as the value of output changed. For this reason the simple relationships drawn in break-even charts no less than in the text-book cost curves can be very misleading.

Selling cost provides a further complication, and also an explanation of why marginal costs might rise even in firms that are capable of producing more goods at constant marginal cost.

The condition of a large output may be heavy selling cost. The curve relating cost of selling and sales might well be the shape of the average cost curve shown in diagram 1. If we add such selling costs to, say, the production costs shown in diagram 2, then aggregate marginal cost might well rise before output K is reached. However, there may not be such a simple relation between cost of selling and sales. On occasion, sales outlay is increased because sales are falling, or because the market is tending to force the firm to produce a costly variety of goods. In some cases, selling cost is used to reduce the product-mix effect on production cost, but as the cost of doing this is likely to vary, it is still necessary to think in terms of the range within which costs move with changes in the value of output.

Inquiry into capacity.

In this section I present some results of an inquiry into the concepts of capacity commonly used in industry and the differences between full capacity, normal capacity and optimum

capacity.[1] The results of this inquiry throw some further light on the relations between cost and output. The following is an account of some typical responses in different industries.

Firm *A*, operating a continuous chemical process, producing a small but variable range of output, distinguishes between *rated capacity*, *boom capacity*, *recession capacity*, and *normal capacity*. The rated capacity is the designed capacity. Designers are frequently over-cautious in their estimates and it is usually found that the plant can be operated safely at 130% of design capacity. This becomes normal capacity.

An increase in the speed of the process lowers quality but as this quality change is not marked, it is safe to push output above normal in time of boom. It is also possible to concentrate on the products that give the most favourable input-output relationships and to turn down orders or lengthen delivery dates for the others. Supplies can also be cut to more competitive overseas markets during conditions of boom at home. Maintenance can also be reduced. For these reasons, boom capacity in terms of value of output at constant prices may rise to 170% of rated capacity.

During recession, process time is reduced below normal to raise quality. This is done to prevent the rejection of deliveries on grounds of unsatisfactory quality. The range of output is also increased, so creating a less favourable input-output relationship. These two reactions may reduce recession output in terms of value at unchanged prices to 80% of normal capacity, i.e. to just above rated capacity. The break-even point is so high that efforts are also made to maintain physical inputs at this level by cutting prices abroad and producing for stock. The firm has several plants and when stocks are thought to have become excessive, a whole plant is closed and not started again until the market is sufficiently good to justify continuous operation at the level of rated capacity.

Full capacity ranges from 80% to 130% of normal, depending on the state of trade. The rise from mild recession

[1]This inquiry was part of a wider inquiry into supply responses. I was ably assisted in this inquiry by Mrs. Cynthia Arditti, who conducted most of the interviews. Our sample was small. It consisted of 40 firms—12 in engineering, 6 in chemicals, 5 in metal processing, 6 in textiles, 4 in building materials, 4 in food and drink, and 3 miscellaneous.

to boom capacity does not increase marginal costs as usually measured but it does reduce quality and consumer choice. Output can be increased smoothly from 80% of normal, but below 80% the change takes place in steps as whole plants are closed.

Firm *B*, undertaking mainly flow but some batch production in chemicals, has a clear-cut concept of capacity. Capacity is defined as continuous operation of plant on 4 shifts. There is standby equipment at some stages because of the need for constant planned maintenance but because this standby equipment is not a complete unit it cannot be used to increase output above normal capacity. Average costs continue to fall as output rises to capacity. The product mix can be varied a little in boom time to increase the value of output at constant prices. The firm has not faced slump conditions, but it supposes that it would react by increasing the range of output and not by cutting the rate of operation. It is not able to estimate the effect of this on the value of output at constant prices.

Firm *C*, which operates both continuous process and normal equipment in fine chemicals, has less definite ideas of capacity. Output from the continuous processes cannot be increased except by the discovery of new methods, but excess demand has induced an intensive study of ways to increase output and, in consequence, output has been raised by a "considerable amount" above rated capacity. Capacity with normal equipment is regarded as output from one shift of 40 hours. Here the reaction to excess demand is to work overtime. If overtime is to be worked for more than a week, they do not work more than 15 extra hours. This would bring output to 115–120% of normal—an increase that implies a marked rise in marginal cost.

When excess demand is maintained the firm introduces a double day shift but as they "try to be permanent employers of male labour (they) work severe overtime until it is clear that (they) can work more permanently." Severe overtime involves "working weekends" as well as longer daily hours.

Capacity is clearly much greater than normal output. I was unable to get a clear explanation of why one shift was regarded as normal. The firm thinks it a nuisance to arrange

the extra supervision, and capacity is planned to meet demand from normal levels of output.

Firm *D*, making a variety of detergents and edible fats by flow and batch processes, uses several concepts of capacity. Their main concept of capacity relates to the potential output of equipment for a given production schedule determined by market conditions. When they say that normally they work at 90% capacity, they mean that "operating inefficiencies" reduce output 10% below maximum for a planned product mix. This could be expressed more clearly in terms used in the iron and steel industry. There is a *planned full load* when the plant is used with maximum efficiency to complete a given production programme. *Normal full load* exists when a given production programme is completed with the usual degree of inefficiency. In this firm, normal full load averages 10% less than planned full load. *Absolute full load* is possible production when all parts of the plant are working continuously at full load. In this firm absolute full load is considerably greater than the normal output. At the time of this survey when we were informed that they were working at 90% capacity, some departments were working 7 days and nights and others 5 eight-hour days. The only sensible answer to the questions about excess capacity would be that with the present product mix there was no excess capacity, but that a more favourable market situation could create excess capacity! It could do this by shifting planned full load towards absolute full load. Because of such complexities it is not possible to give simple statistics of the movement of cost with output. For individual departments the movement of cost is similar to that shown in the diagram of Table II.

Firm *E*, in brewing, uses a concept of normal capacity for standard costing purposes. This is 85% of full capacity.[1] Full capacity output has, however, been increased from time to time. When after the war the annual increase in demand continued at an unexpectedly high level, efforts were made to increase output from existing plant. At that time, engineers

[1]As the firm does not operate at a limited peak and has a standard product mix, its concept of capacity is not ambiguous as alleged by Vaisey for brewers in general. *v. The Brewing Industry*, Chapter 6.

shut down plant 2 days a week for maintenance and cleaning, but faced with the demand for more output they were stimulated to realize that "the easiest way to keep vessels clean is to keep them full" and the plant is now shut down only once a fortnight. This increased full capacity by 30%. Normal capacity is little short of full capacity allowing for this shutdown. Full capacity is also optimum capacity in the sense that this is the lowest cost situation.

Where, as with many brewers, the product mix is not standard, capacity may vary in the manner already outlined for Firm *A*.

Firm *F*, a mass producer of food products, uses much the same concepts of capacity as firm *D*. There is a similar complication in applying a precise concept of capacity.

Firm *G*, which uses a one-shift mixed flow and batch process for the preparation of leather, thinks of its capacity as x square feet of leather per year. This concept of capacity is however derived from the present market situation. They could prepare 50% more at the wet end of the process, but because leather is now finished in small batches (partly to meet the demands of the fashion trade for fancy colours,) the finishing end could not handle the additional output. The firm is "thinking of adding to the finishing department" but has doubts about the future of the industry.

Because of the difficulty of getting labour for work that is unpleasant and dirty at many stages, shift working is not possible. Output is occasionally raised above normal by some evening and Saturday work and they can increase the size of batches and so create room for greater output from the wet end. It was not possible to get information on the relation of full capacity in this sense to normal capacity. The reaction to a slump in demand is to increase the product range and to put workers on short time.

Firm *H* makes paper by continuous and semi-continuous methods. It runs one modern plant continuously (i.e. on a four shift basis) apart from a fortnight in summer, and one older plant for 5½ days (3 shifts) except in times of sufficient excess demand when the crews "work weekends." The concept of capacity used is that of the amount of production needed to

keep these plants going for 7 and 5½ days respectively. The movement of cost in the three-shift mill is very similar to that shown in the diagram of Table II, with marginal cost rising sharply for "over capacity" or weekend working. This diagram could easily be extended to provide a picture of costs for four-shift working.

This total-hours concept of capacity can be confusing. Plants can make different grades of paper and the value of output (at constant prices) can move through a considerable range. Thus, when demand falls it is not unusual to shift to cheaper grades of paper to maintain capacity working. Where it is not possible to maintain capacity the days of working are reduced, but the paper produced may or may not be normal.

Firm *J* makes bricks by semi-continuous production methods. It has a definite concept of capacity—the maximum possible number of bricks that can be produced with its equipment. X men working "5½ days and normal overtime" produce enough to maintain drying and firing equipment in full use. Average costs fall to this level of output. As the firm makes a standard brick, differences between capacity and normal output are due to interruptions caused by illness, weather, or breakdown of machines.

Firm *K*, making pottery by mixed flow and batch processes, works to the capacity set by the continuous tunnel oven. It was said that output tends to be 90% of capacity. This calculation however, makes no allowance for possible variations in the quality of decorated ware. It would be possible to increase the quality of decoration, partly by employing more decorators, and so raise the value of output without adding to oven capacity. The degree of such variation in capacity output in short periods is relatively small.

It is also possible to increase speeds through ovens. However, without change in ancilliary equipment, this would require overtime working. As wage cost is a large part of total cost in ancilliary processes, the rise in wage rates brings a significant increase in variable costs per unit. This makes overtime working unprofitable except when excess demand is sufficient to push up pottery prices relative to wage rates.

Firm *L*, an integrated steel works, distinguishes between absolute and normal capacity, for reasons given when discussing firm *D*. There is now permanent excess capacity in certain parts of the works due to technical change. For example, it has excess capacity in coke ovens since the installation of coal-blending and sintering plants that have reduced the weight of coke per ton of pig iron. The firm also chose to instal a rolling mill with capacity greater than the output of the melting shop, and a blast-furnace plant below the largest possible requirements of the melting shop on the grounds that cost factors make it imperative to keep furnaces going to the full. Because of the range and variability of output the firm does not use the concept of optimum capacity. Unfortunately, the firm was not able to give figures for the range of output value when working at *capacity*.

Firm *M*, engaged in the batch production of engines, treats its capacity as normal output from two shifts. However, this normal is an average—the product mix is such that sometimes they exceed it and sometimes fall below it. To a certain extent they "can influence the mix by price concessions, but optimum utilization can't be reached because of the variety of the product." Subcontracting is standard practice, but output is usually limited by the capacity of certain large and expensive machines producing parts that cannot be subcontracted. This apart, flexibility in output is provided by overtime and subcontracting.

Variable costs per unit change little with output, labour cost per unit is low, and average cost falls with the degree of machine utilization.

Firm *N*, making high quality machinery for a range of industries, does not use a concept of capacity, though they "like to keep plant fully employed." Yet, when thinking of orders they are never guided by manufacturing capacity, only their ability to make a good machine. Their capacity is set in a general way by the type of machine ordered, the supply of competent design staff and the capacity of subcontractors.

Firm *O*, engaged in spinning, reckons its capacity in terms of spindle hours. However, as it was in the process of shifting its normal hours from 45 to 75, this capacity measure was rather

confusing. It also uses a concept of balanced capacity. For example, if the carding machines were used for a coarse yarn, all the carding machines would be loaded but the spinning machines would be half idle. If, on the other hand, a very fine yarn was produced the card room would have excess capacity. In this firm, capacity in the card room and spinning room are in balance at a 35 count, or when production is split between two extremes such as 80 and 12.

Costs were said to move in much the same way as those shown in Table 1. It is, however, clear from their concept of balanced capacity that this is only so for a balanced output. Unfortunately the firm was not able to give information on the range of costs that might apply to a given level of output value. It follows, of course, that when times are sufficiently good, the firm can, without changing prices, increase the value of output from a given labour input by choosing the most favourable product mix.

It is clear from this short treatment of capacity in these fourteen firms that it is often necessary to distinguish between several concepts of capacity. Measures of capacity in use may be required for purposes of production control, for budgeting, for calculating costs, for fixing prices or for estimating delivery times. The appropriate concept of capacity often varies with the purpose of the measure. When we are wondering whether an expansion in effective demand would be inflationary or estimating the safe amount by which effective demands could be expanded, it is not appropriate simply to ask firms whether they could increase their output and if so by how much. In multi-product firms there is no one level of capacity. Capacity varies with the nature of demand.

The cost of expansion also varies with the nature of the expansion depending on whether hours are extended by ending short-time working, by working overtime, by adding a shift, or by adding to the labour force in a given shift. Ending short time working and adding (efficient) labour to run (efficient) laid off equipment are the cheapest ways to increase output. In such cases, an expansion of output will take place at lower average costs per unit so long as materials and bought-in components are available. In my sample of firms there were

few cases of plant made up of several or many similar machines (as with a cotton mill) and in most cases labour inputs were increased by ending short-time or working overtime or adding another shift.

The effects of overtime and shift working vary considerably. In several firms (e.g. in the pottery, bricks, and engines firms mentioned) normal capacity is set by one piece of plant and overtime is only called for when the product varies from normal or when interruptions to work flow have to be made good. In pottery, overtime working in many of the processes adds so much to cost that even when labour is short, overtime is used only when product prices rise relative to wage rates. In several cases mentioned, overtime-output is more expensive than another-shift output, but whether overtime or another shift is chosen depends on whether labour is available for shift work, on the extent of excess demand, and on forecasts of the trend of demand. There is not a smooth movement from overtime to second (or third) shift, but a substantial jump. Considerable excess demand may therefore lead to expansion of output at lower marginal cost than slight excess demand. (cf. the diagram to Table II).

The value of output at capacity is also capable of varying significantly without a change in commodity prices. The most striking case of this was provided by firm *A* (in chemicals) where the variation is large. Smaller but significant variations were possible for firms *D* (detergents and edible fats), *F* (food), *G* (leather), *H* (paper), *K* (pottery), *L* (iron and steel), *M* (engines), and *O* (cotton spinning). In all these cases an improvement in their market situations could produce a very marked increase in productivity as it is usually measured.

In many cases marginal cost for such an expansion of output remains constant (at given factor prices). However, not all increases in effective demand will have this effect on firms with excess capacity loosely defined. The increase in effective demand may not "improve the market situation" in the appropriate way. Increased value of output may thus take place at increased real cost as firms meet increased, but not appropriately balanced, demand by more overtime working not accompanied by a closing of the gap between absolute full

load and planned full load (as defined for firm *D*). This failure to close the gap may be due to the nature of demand or to failure or inability of suppliers to provide an appropriate flow of materials and components of the appropriate quantities or qualities.

If real cost as defined does not rise, it does not follow that there will not be inflation. The rise in effective demand may force up production costs through inducing a rise in factor prices. Whether this will happen will depend on such things as the trend of investment in labour-saving equipment (which will determine the precise relation between the demand for goods and labour), and on institutional arrangements for determining salaries and wages.

BRUCE WILLIAMS

University of Manchester.